*Volume 2* *May, 1932* *Number 5*

# PHYSICS

## Dielectric Properties of Some Glycols

*By* A. H. WHITE *and* S. O. MORGAN
*Bell Telephone Laboratories*

(Received March 4, 1932)

The Debye theory of polar molecules leads to the conclusion that the anomalous dispersion due to dipoles which has been observed at high frequencies should also be found at low frequencies. Supercooled glycols are viscous enough at low temperatures to show dispersion at low frequencies. Real and imaginary dielectric constant and power factor are calculated and their inter-relationships are discussed. Better than qualitative agreement with the Debye theory is observed in the pure pentandiol, but not in the solution measurement.

IT IS well known that in applying the Debye equations for determining the electric moment, the results obtained for pure liquids, especially those ordinarily considered to be associated, do not agree with those found for dilute solutions or gases. Although the simplifying assumptions of the Debye theory are no longer to be regarded as valid for associated liquids or for solids, such a treatment appears qualitatively correct and is useful as a means of explaining the effect of temperature and frequency upon dielectric properties. For the purpose of relating dielectric properties to chemical and physical structure, measurements on pure liquids are quite as important as dilute solution measurements.

A study of the anomalous dispersion, i.e., change of dielectric constant and dielectric absorption with frequency, provides a useful test of the Debye theory. This change of dielectric constant with frequency in the electric spectrum is usually called anomalous dispersion to distinguish it from the normal dispersion observed at optical frequencies. Anomalous dispersion is observed for highly damped polarizations, as for example for dipole polarization, where the time of relaxation is determined primarily by viscous forces. Normal dispersion is found for lightly damped polarizations such as electronic polarization where the inertia of the charges is the determining factor.

Most of the measurements of the variation of dielectric constant with frequency which have been made on pure materials have been made at high frequencies. Mizushima's[1] measurements on a series of alcohols, for example, were at 0.57, 3, 9, and 50 meters. Johnstone and Williams[2] measured solu-

[1] S. Mizushima, Bull. Chem. Soc. Japan **1**, 47, 83, 115, 143, 163 (1926).
[2] J. H. L. Johnstone and J. W. Williams, Phys. Rev. **34**, 1483 (1929).

tions of nitrobenzene and para-dichlorobenzene in a viscous oil at several frequencies between 34 and 154 meters. Goldammer and Sack,[3] working at 52.8, 77.5 and 139 cm, showed a change of dielectric constant with frequency for solutions of some alcohols in a mineral oil. Of these investigators only Mizushima has given conductance data from which the absorption accompanying the anomalous dispersion can be determined. There have been numerous measurements[4] on rosin, insulating oils and mixtures of the two, but the structure of these materials has usually not been known.

TABLE I.

| Substance | B.P. | $N_D^{20°}$ |
|---|---|---|
| Propylene glycol | 65°–66° at 9 mm | 1.4323 |
| Trimethylene glycol | 87°–88° at 9 mm | 1.4396 |
| 2-Methyl pentandiol 2.4 | 92°–93° at 7 mm | 1.4316 |

There is no reason from the theory for believing that the same anomalous dispersion shown for example by Mizushima for the alcohols, would not also be found at low frequencies. Most materials do not show this dispersion at

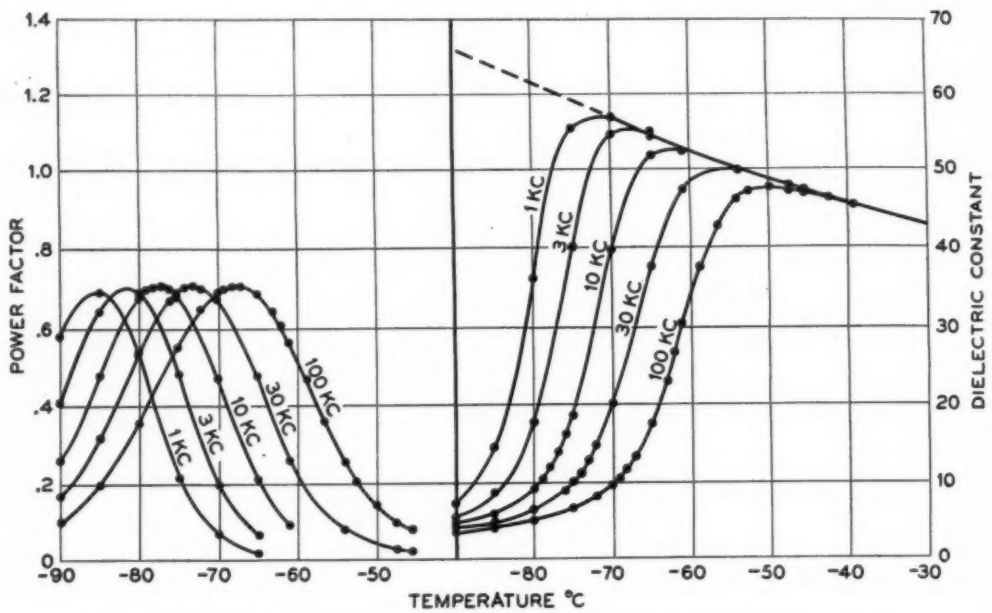



Fig. 1. Dielectric constant and power factor vs. temperature for frequencies from 1 to 100 kilocycles for propylene glycol.

low frequencies due to the fact that their viscosity is not sufficiently high. However, such polar liquids as may be supercooled to form glasses usually

[3] R. Goldammer and H. Sack, Phys. Zeits. **31,** 224 (1930).

[4] D. W. Kitchen and H. Mueller, Phys. Rev. **32,** 979 (1928); E. Kirch and W. Riebel, Arch. f. Electrot. **24,** 355, 553 (1930); L. S. Ornstein and G. Willemense, Zeits. f. tech. Physik **11,** 345 (1930); H. H. Race, Phys. Rev. **37,** 430 (1931).

do show anomalous dispersion. The glycols provide a series of substances which become sufficiently viscous at low temperatures to show this dispersion at low frequencies and this paper describes the results of measurements on three glycols. The characteristics of the glycols used are given in Table I. These materials were dried and then distilled at low pressure.

The measurements were made with a capacitance bridge described by Shackelton and Ferguson[5] and were direct capacitance and conductance measurements. They were usually made with a descending order of temperatures but sufficient time was allowed after changing the temperature for the dielectric constant and conductance to become constant. That this was so is indicated by the fact that points taken with ascending order of temperatures fall on the same curves as those taken with descending temperatures.

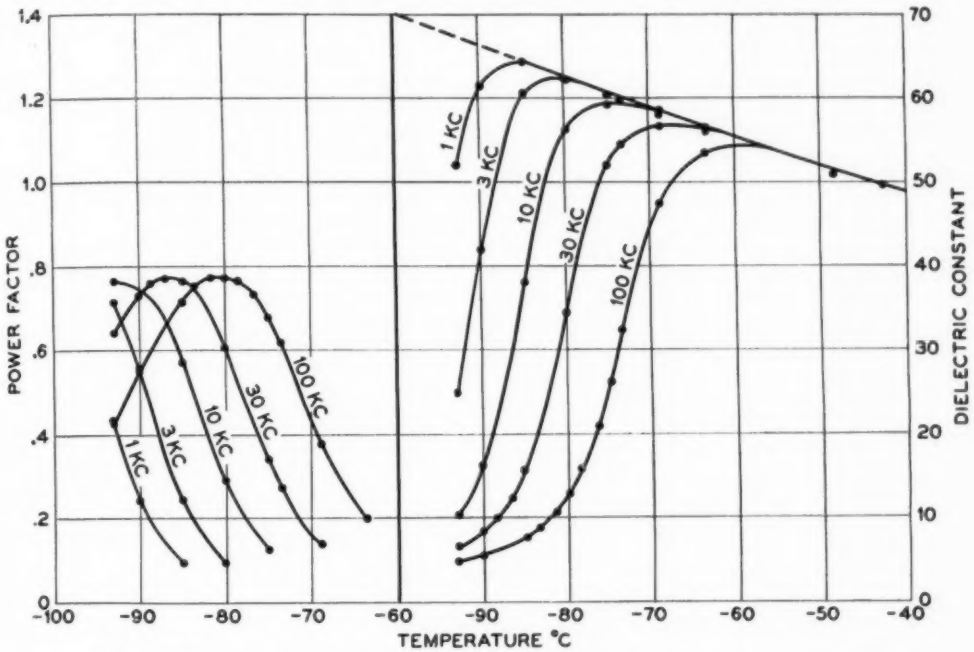



Fig. 2. Dielectric constant and power factor vs. temperature for frequencies from 1 to 100 kilocycles for trimethylene glycol.

In Figs. 1 to 4 are shown curves of dielectric constant and power factor for the three glycols and the solution of pentandiol in the non-polar solvent dioxane. The points are the observed values. It was not possible to observe any dispersion for very dilute solutions of these glycols although, as can be seen from Fig. 4, essentially the same behavior is observed for a concentrated solution as for the pure glycol, it merely being displaced to a lower temperature as required by the theory. Bridge measurements can not be made at frequencies such as would be necessary to observe dispersion for the dilute solutions and high frequency conductance measurements by other methods are more difficult to make and less dependable.

[5] W. J. Shackelton and J. G. Ferguson, Bell System Tech. J. **7**, 70 (1928).

It is apparent from the curves that the different glycols, and even the solution, all show the same general behavior. The curves are, however, dis-

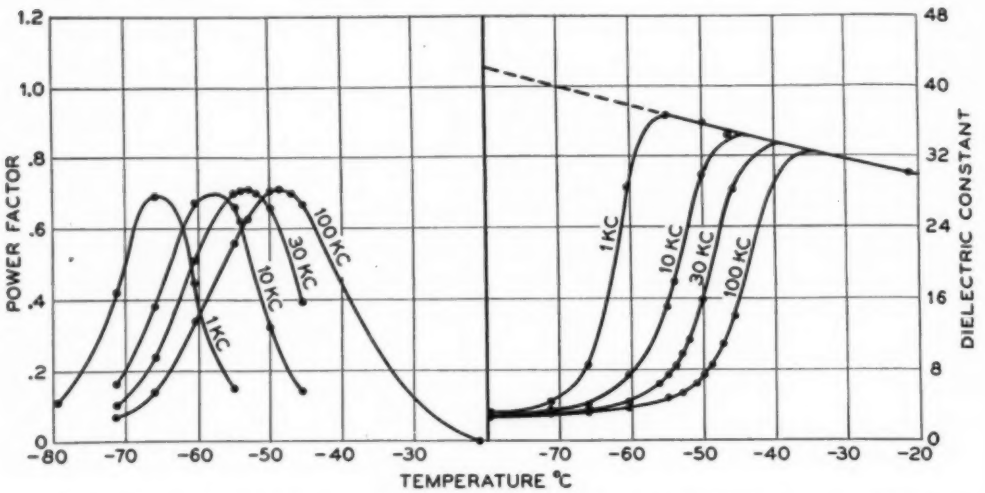



Fig. 3. Dielectric constant and power factor vs. temperature for frequencies from 1 to 100 kilocycles for 2-methyl pentandiol 2.4.

placed along the temperature axis, the dispersion of dielectric constant first appearing for pentandiol, Fig. 3, at $-35°$ while for trimethylene glycol, Fig.

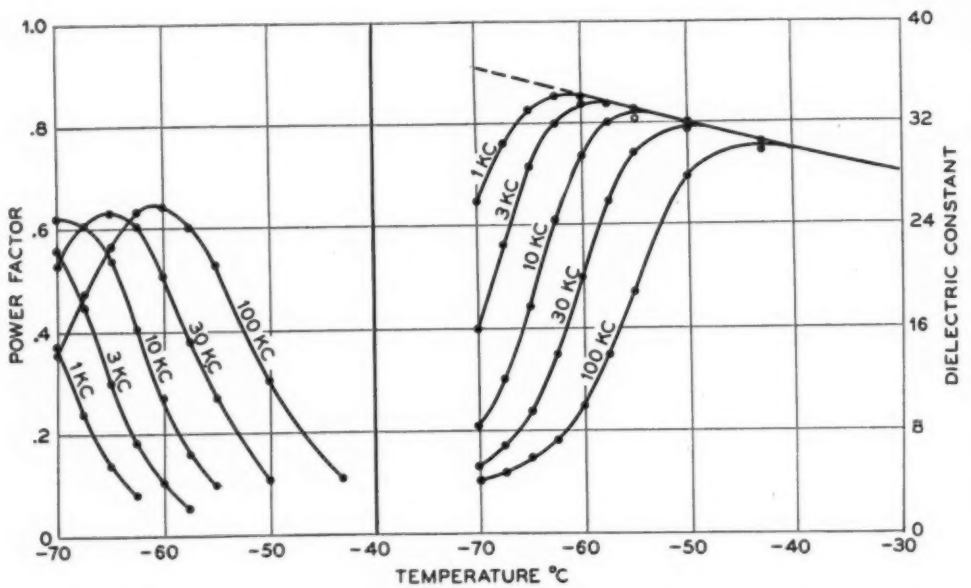



Fig. 4. Dielectric constant and power factor vs. temperature for frequencies from 1 to 100 kilocycles for 0.83 MF solution of 2-methyl pentandiol 2.4 in dioxane.

2, it starts at $-60°$. At the higher temperatures the dielectric constant of each of these materials is independent of frequency in the range below 100

kilocycles and is increasing rapidly with decreasing temperature. As the temperature is lowered the viscosity increases and the time of relaxation of the dipole, which is given by the equation

$$\tau = \frac{4\pi a^3 \eta}{kT}, \tag{1}$$

where $\eta$ is the viscosity, $a$ the molecular radius and $T$ the absolute temperature, increases. When this becomes of the same order as the period of the applied field a decrease of dielectric constant is observed. At this temperature also the power factor starts to increase and it reaches a maximum near the temperature where the change of dielectric constant with frequency is greatest.

The general equation for the variation of dielectric constant with frequency as given by Debye is

$$\epsilon = \frac{\dfrac{\epsilon_0}{(\epsilon_0 + 2)} + i\omega\tau \dfrac{\epsilon_\infty}{(\epsilon_\infty + 2)}}{\dfrac{1}{(\epsilon_0 + 2)} + i\omega\tau \dfrac{1}{(\epsilon_\infty + 2)}}$$

where $\epsilon_0$ is the static dielectric constant or the value at a low enough frequency that dipole orientation is constant, and $\epsilon_\infty$ is the high-frequency dielectric constant where the effect of dipole orientation has disappeared. The equation for the dielectric constant may also be written $\epsilon = \epsilon' - i\epsilon''$. The equations for $\epsilon'$ and $\epsilon''$ in terms of $\epsilon_0$, $\epsilon_\infty$ and $\tau$ are

$$\epsilon' = \epsilon_\infty + \frac{\epsilon_0 - \epsilon_\infty}{1 + \left[\dfrac{\epsilon_0 + 2}{\epsilon_\infty + 2}\right]^2 \omega^2\tau^2}$$

$$\epsilon'' = (\epsilon_0 - \epsilon_\infty) \frac{\dfrac{\epsilon_0 + 2}{\epsilon_\infty + 2}\omega\tau}{1 + \left[\dfrac{\epsilon_0 + 2}{\epsilon_\infty + 2}\right]^2 \omega^2\tau^2}.$$

For the case of a condenser shunted by a resistance, $\epsilon' = C/C_0$ and $\epsilon'' = G/\omega C_0$, $C$ being the total capacity, $C_0$ the capacity of the empty condenser and $G$ the equivalent parallel conductance. This conductance is the non-d.c. conductance. The term $\epsilon''$ which is a measure of the dielectric absorption should be a maximum when

$$\omega\tau = \frac{\epsilon_\infty + 2}{\epsilon_0 + 2}.$$

The curves in Figs. 1 to 4 show the maxima for each frequency and also show that this is displaced toward lower temperatures (greater viscosity) for low

frequencies. From the position of these maxima and the values of $\epsilon_\infty$ and $\epsilon_0$, $\tau$ may be determined and thus the general curve of variation of dielectric constant with frequency calculated. This has been done for all of the pure glycols and for the solution, and very good agreement between calculated and observed values is found. Table II gives the values of $\epsilon_\infty$, $\epsilon_0$ and $\tau$ used in calculating the theoretical curves.

TABLE II.

| Substance | Temp. | $\epsilon_0$ | $\epsilon_\infty$ | $\tau$ |
|---|---|---|---|---|
| Propylene glycol | −75° | 59.5 | 3.5 | $2.8\times10^{-6}$ |
| Trimethylene glycol | −90° | 66.4 | 3.2 | $3.0\times10^{-6}$ |
| Pentandiol | −60.4 | 39 | 2.7 | $1.0\times10^{-5}$ |
| 0.83 MF solution of pentandiol in dioxane | −65 | 35.1 | 2.6 | $2.4\times10^{-6}$ |

In Fig. 5 the solid curves of $\epsilon'$ and $\epsilon''$ are calculated from the theory for the pentandiol at −60.4°C and the points are the observed values. The agreement is very good. Fig. 6 shows similar caluclated curves and experimental points for the 0.83 MF solution of pentandiol in dioxane, and the agreement

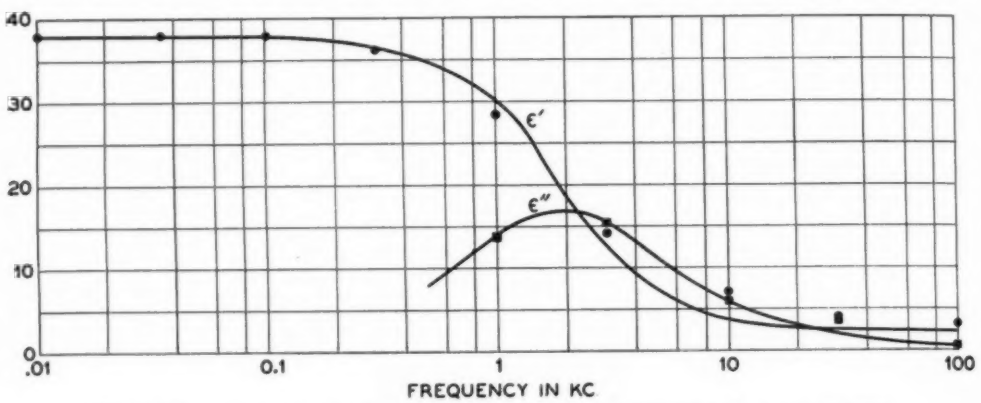



Fig. 5. Curves of $\epsilon'$ and $\epsilon''$ vs. frequency for 2-methyl pentandiol 2,4 at -60.4°C.

is only qualitative. For the propylene and trimethylene glycols the agreement is also only qualitative but it is good enough to regard the Debye theory as being very useful in explaining the dielectric behavior of polar liquids.

The case of the solution of pentandiol in dioxane, a nonpolar solvent, is particularly worth comment since it is to be expected that the best agreement between theory and experiment would be found for dilute solutions, where the assumptions of the Debye theory are most nearly applicable. The agreement here is the poorest of any. This solution, however, is not in any sense a dilute solution and the peculiarities of the curves of polarization vs. concentration for alcohols as shown by Smyth and Stoops[6] and for glycols as will be shown in a later paper show that complex associations or compound formations are taking place in such solutions so that the observed time of relaxation is probably not that of a single molecule but an average value of

[6] C. P. Smyth and W. N. Stoops, J.A.C.S. **51**, 3312 (1929).

that for several combinations of molecules each having different times of relaxation.

In Table III are given the values of the electric moments as determined from dilute solution measurements in dioxane and in one case in n-heptane. The third and fourth columns respectively give the values of the maxima of power factor and $\epsilon''$ for the pure glycols and the solution of pentandiol in

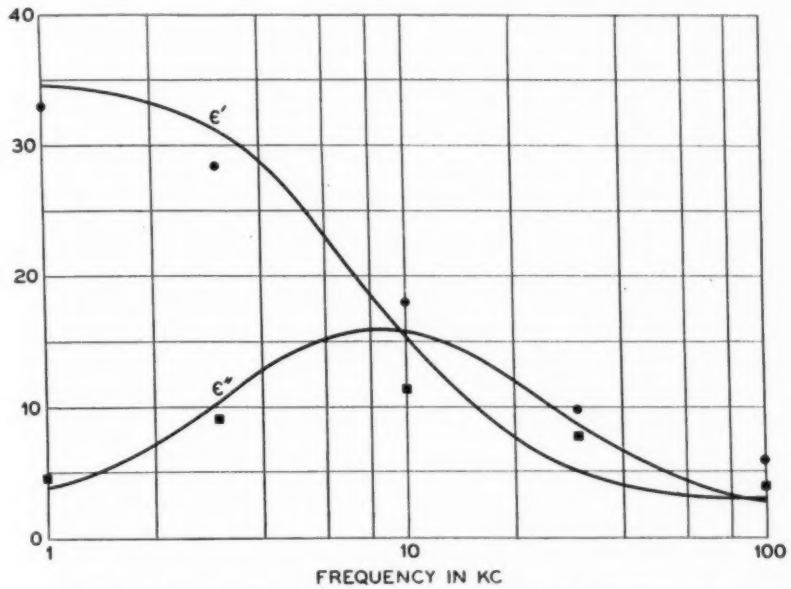



Fig. 6. Curves of $\epsilon'$ and $\epsilon''$ vs. frequency for 0.83 MF solution of 2-methyl pentandiol 2.4 in dioxane at $-65°$C.

dioxane and the fifth the dielectric constant at $-30°$. There is no quantitative agreement between the values of dielectric constant or power factor of the pure glycols and the electric moment. As was pointed out earlier, it has been found, quite generally, to be impossible to deduce the properties of associated liquids from dilute solution studies.

TABLE III.

| Substance | $\mu \times 10^{18}$ | Power factor max. | $\epsilon''_{max.}$ | $\epsilon_{-30°}$ |
|---|---|---|---|---|
| Propylene glycol | 2.2 | 0.71 | 24 | 43 |
| Trimethylene glycol | 2.35 | 0.77 | 28 | 46 |
| 2-Methyl pentandiol 2.4 (dioxane) | 2.9 | 0.71 | 17 | 31.7 |
| 2-Methyl pentandiol (*n*-heptane) | 2.1 | | | |
| 0.83 M.F. solution pentandiol | | 0.64 | 12 | 28 |

The values of the power factor maxima for the pure glycols are not very different from each other and may represent the range which is characteristic of glycols. The power factor as calculated here is $G/(G^2+\omega^2C^2)^{1/2}$ where $G$ is the equivalent parallel conductance. In all cases the direct current conductance had decreased at temperatures where absorption was observed until

it amounted to only a few percent of the total for frequencies above $10^3$ cycles. The power factor is the ratio of the in-phase current to the total current and is not directly significant in connection with the Debye theory. What is more significant is the value of $\epsilon''$, the imaginary part of the complex dielectric constant. It is apparent from Table III that the differences in the value of the total dielectric constant (fifth column) are more nearly reflected in the value of $\epsilon''$ than of power factor.

As has been pointed out before there is some question as to the applicability of the Debye equations to such highly associated substances as the alcohols and glycols. Although the limitations of the theory may preclude its quantitative application and its use for the calculation of molecular constants from data on pure liquids, still it seems that qualitatively at least the rotation of dipoles may be offered as a valid explanation of the anomalous dispersion and absorption in substances known to contain polar molecules.

There is the possibility that this anomalous dispersion and absorption which has been observed for alcohols, glycols and other polar substances may be due to an ionic polarization rather than to dipoles. A brief discussion of this possibility is given in a paper by the authors[7] reporting similar measurements for rosin oils. The ionic polarization is an especially attractive explanation only at low frequencies. It can hardly apply at high radio frequencies. Our results, as well as those of others show however that the value of the maximum of absorption is essentially the same at low frequencies as at high although our range of frequencies is not high enough to exclude definitely the possibility of ionic polarization. This is the behavior predicted by the Debye theory but is not what would be expected if ionic polarization were responsible for the anomalous dispersion at low frequencies and dipole polarization at high. The constant value of the maxima of the absorption from $10^3$ to $10^5$ cycles in the case of the glycols is strong evidence that the cause is the same at all frequencies namely the rotation of dipoles.

One further test of the applicability of the Debye theory to these results is found in the comparison of the time of relaxation as obtained from the dielectric data with the value calculated from molecular radius and viscosity by Eq. (1). Alternatively the molecular radius may be calculated from the di-

TABLE IV. *Viscosity and calculated molecular radius for 2-methyl pentandiol 2.4.*

| Temp. | $\eta$ (Poises) | $T\tau$ | $a$ (cm) |
|---|---|---|---|
| −45.4 | $1.77\times10^{3}$ | $6.6\times10^{-3}$ | $3.4\times10^{-8}$ |
| −49.8 | $5.75\times10^{3}$ | $1.8\times10^{-3}$ | $3.2\times10^{-8}$ |
| −54.9 | $2.5\times10^{4}$ | $5.8\times10^{-2}$ | $3.0\times10^{-8}$ |

electric data. Such calculations are to be considered as interesting indications rather than proof of the correctness of the theory. A discussion of the limitations involved in the use of the ordinary coefficient of viscosity for such calculations is given by Williams and Oncley.[8] In Table IV are given the values of

[7] S. O. Morgan and A. H. White, J. Franklin Inst.

[8] J. W. Williams and J. L. Oncley, Jour. of Rheology **2**, 271 (1931).

viscosity and calculated molecular radius for the pentandiol at three low temperatures in the range where dispersion was observed.

Such calculations are not considered by the authors to be more than an approximation but it appears significant that the results are of the correct order of magnitude. Mizushima's calculation of the molecular radius for the alcohols also agreed well with the values obtained by other means but his value for glycerine was too low. With the exception of glycerine the calculations of molecular radius from dielectric dispersion data for pure substances have given results of the right order, which should be considered more than fortuitous.

The time of relaxation of the dipole has been related to the viscosity and molecular radius by Debye as given in Eq. (1). This equation should not be expected to be rigidly applicable for associated substances such as the glycols. Williams and Oncley[8] have pointed out the difficulties involved in the use of the ordinary coefficient of viscosity for calculating molecular radius from dispersion data. We have measured the viscosity of pentandiol in the range of temperatures where dispersion was observed and from these data and values of $\tau$, determined from the positions of the absorption maxima, have calculated its molecular radius. The result is a value of $0.75\times10^{-8}$ cm. This is too low but the result is not unexpected. It is interesting to note that the value falls between that of $0.35\times10^{-8}$, calculated by Mizushima from similar data for glycerine, and $2.0\times10^{-8}$, his value for n-propyl alcohol.

# Development of Polar Characteristics in Insulating Oils*

By W. N. Stoops
*Research Laboratories, Westinghouse Electric and Manufacturing Company*

(Received March 7, 1932)

It is shown that the decomposition in transformer oils may be followed rather closely by measuring the dielectric constant of the oils as they deteriorate in service, the method being sufficiently accurate to distinguish between different oils. These data are compared with those obtained on oils decomposed very much more rapidly in laboratory tests. The change in polar characteristics is definitely correlated with change in chemical characteristics. That the change in dielectric constant of the oils is due to the formation of polar molecules is shown by a study of the variation of the dielectric constant and power factor of the decomposed oils with temperature and frequency.

## Introduction

THE oil used in oil-filled transformers and circuit breakers is a highly refined product containing little or no unsaturated material. Saturated hydrocarbons have been shown to be nonpolar,[1] so that insulating oils are practically non-polar, that is, the molecules are electrically symmetrical. However, in use such oil gradually decomposes, largely due to oxidation, and forms various oxidation products, chiefly acids, esters, and some solid material or sludge, eventually becoming unfit for use. These decomposition products are all polar. Hence it appeared likely that a study of the change in properties of the oil such as dielectric constant and power factor that depend on polarity might be useful in following the course of the decomposition as it occurs in a transformer. The following results were obtained in a study of this matter.

## Experimental Method

The dielectric constant measurements were made with a capacity bridge previously described.[2] A change in dielectric constant of one part in five hundred was easily detectable. The measurements were all made at 300,000 cycles.

A resonance circuit recommended by the American Society for Testing Materials[3] was used to make the power factor measurements. It was modified by introducing an intermediate tuning circuit between the oscillator and the measuring circuit to remove any harmonics in the wave form. The same cell and method of temperature control[2] were used for both dielectric constant and power factor measurements. The power factor tests were made at two frequencies, 500,000 and 1,000,000 cycles.

* Scientific paper No. 581.
[1] Smyth and Stoops, J.A.C.S. **50**, 1883 (1928).
[2] Stoops, J. Phys. Chem. **35**, 1704 (1931).
[3] Proc. A.S.T.M. Vol. 31 I. 877, and 1931 T.S., 642.

The three oils studied were first oxidized by heating a small sample of each in an open bottle in an oven kept at 160°C. The dielectric constant of each oil was measured every day or two for the duration of the test. For this purpose a 20 cc portion was withdrawn, its dielectric constant measured at 30.0° C and then returned to the remainder of the oil in the oven. To determine the importance of the oxygen in the air in the development of polar characteristics in the oils, glass bulbs were secured in which an atmosphere of nitrogen could be maintained above the oil, and from which samples could be withdrawn for testing without introducing any air. The oils were as before all heated at 160°C. Metallic copper is known to be an effective catalyst in the decomposition of transformer oil. Accordingly two samples were heated under the same conditions obtaining in the first set of experiments, except that each bottle contained about two feet of bare copper wire. One of these samples was kept closed and the other left open during the heating.

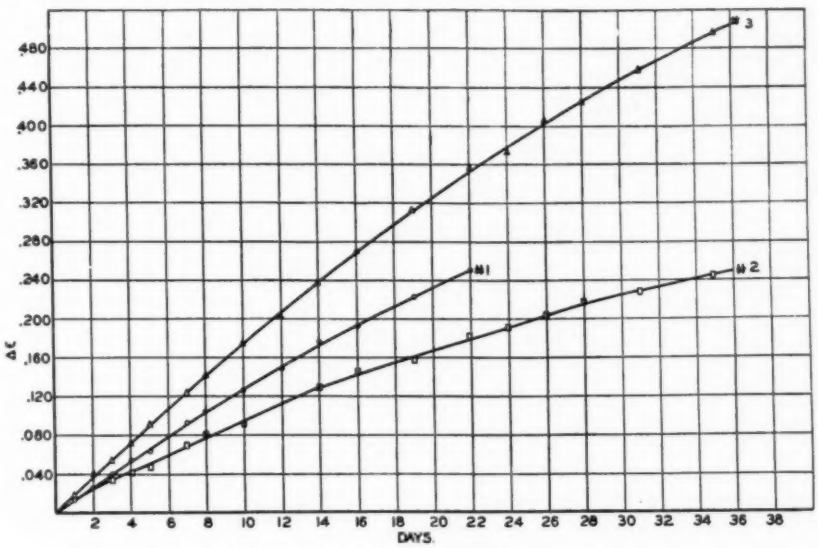



Fig. 1. Oils heated in air.

## Materials

The oils used in these tests were of different compositions. Oil No. 1 was prepared from a mid-continent crude and consisted largely of naphthenic hydrocarbons with about 20 percent of aromatic hydrocarbons and a very small amount of resinous material. No. 2 was a more highly refined oil from the same crude, being almost 100 percent naphthenic hydrocarbons. Oil No. 3 was prepared from a Pennsylvania paraffin base crude, and was composed almost entirely of paraffin hydrocarbons with no unsaturated, aromatic or resinous compounds.

These three oils were chosen for study because of their different compositions and because they had been used in transformer life tests. That is, their decomposition in service had been studied, samples of oil being drawn from the transformers every few months for testing. Dielectric constant measurements were also made on these samples.

A sample of oil No. 2, oxidized in the laboratory, was used in the first series of power factor measurements. This oil solidifies between −50°C and −60°C. Another oil which solidifies at about −10°C was oxidized and used in further experiments. In order to increase its viscosity, it was mixed with an equal quantity of a pure hydrocarbon lubricating oil of viscosity at 40°C, 18 times greater and of about the same freezing point.

## Results

The data are all shown graphically. In Figs. 1, 2 and 3, the increase in dielectric constant of the oils is plotted against the time of heating. In Figs. 1 and 2, curve No. 1 represents oil No. 1, curve No. 2 oil No. 2, etc. Similar

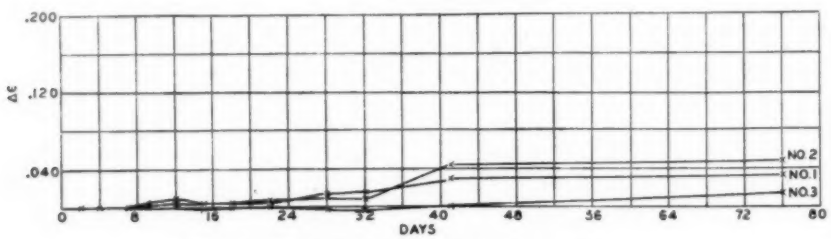



Fig. 2. Oils heated in atmosphere of nitrogen.

data for the same oils, but obtained from measurements on samples withdrawn from transformers in service, are shown in Fig. 4. The data obtained in one of the chemical tests on these samples are shown for comparison in Fig.

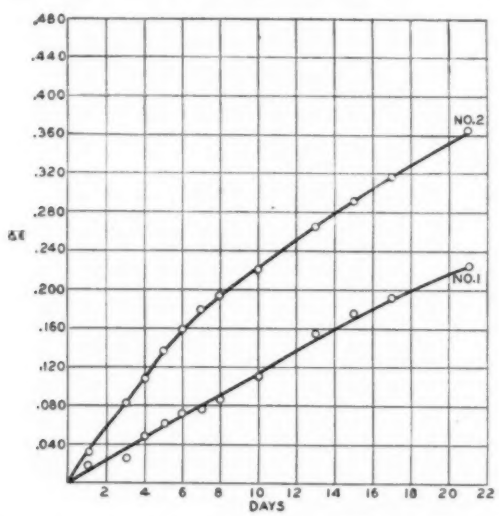



Fig. 3. Oils heated in air in presence of copper. No. 1:—Oil No. 1 heated with limited access to air. No. 2:—Oil No. 1 heated with free access to air.

5. Figs. 6, 7 and 8 show the variation of dielectric constant and power factor with temperature for these oils. The unused oil of Fig. 6 is very similar to oil No. 1.

## Discussion of Results

It may be pointed out that the dielectric constants of these oils when new give some indication of their composition. Aromatic hydrocarbons are known to have larger dielectric constants than naphthenic hydrocarbons, which in turn are larger than those of the paraffin hydrocarbons. For instance, the dielectric constant of benzene, an aromatic hydrocarbon, at 30°C is $2.26_4$; of cyclohexane, the corresponding naphthenic hydrocarbon, $2.01_3$, and of hexane, the corresponding paraffin hydrocarbon, $1.87_5$. The dielectric constants of the new oils are 2.241, 2.159 and 2.110 for No. 1, 2 and 3 respectively. Oil No. 1 containing an appreciable percentage of aromatic compounds, has the largest dielectric constant, and No. 3, the paraffin base oil, has the smallest value.

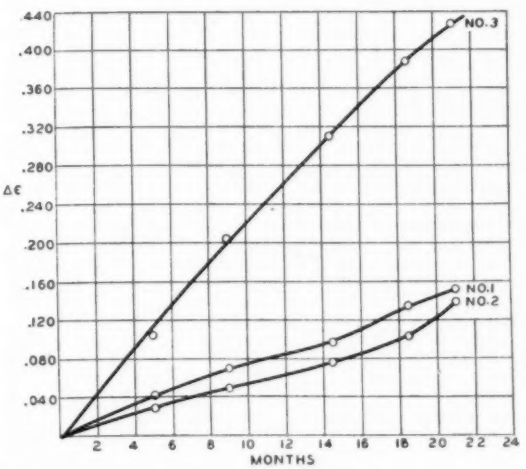



Fig. 4. Change of dielectric constant of oils in use in transformers.

The oils show a marked difference in the development of polar characteristics. The final oxidation products under these conditions are chiefly acids and esters. The formation of a certain amount of solid matter or sludge also occurs. The oxidation products are all polar and have relatively large dielectric constants. Of course, as these compounds are formed in the oil, the dielectric constant of the mixture increases. That the presence of the oxygen in the air was the most important factor in the destruction of transformer oils in service is well known, and is brought out clearly here by the striking difference in the development of polarity in the oils when they are heated in air and when under an atmosphere of nitrogen, as shown in Figs. 1 and 2. The catalytic influence of copper is very clearly shown in Fig. 3. Oil No. 1 was used in this test and hence curve No. 2 is comparable with curve No. 1 of Fig. 1. It is readily apparent that the presence of the copper almost doubles the rate of decomposition. The change in the dielectric constants of the oils as they develop polar characteristics in service is shown in Fig. 4. These curves are very similar to those of Fig. 1, and indicate that the results obtained in the laboratory tests closely parallel those obtained in service. The polar substances

formed by heating oil exposed to air are practically all oxidation products, chiefly acids and esters. A chemical test which measures the amount of acids and esters formed should show at least rough parallelism to the change in dielectric constant. Such data are shown in Fig. 5 and the similarity to Figs. 1 and 4 is apparent. A further proof of the polar nature of the used oils is shown

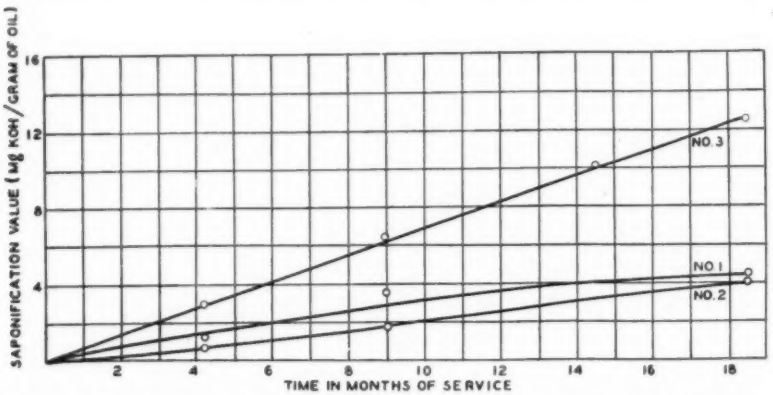



Fig. 5. Development of spaonifiable matter in transformer oils with service.

in Fig. 6. Here the dielectric constants of the oils are shown plotted against temperature. Their dielectric constants do not differ much at the extremely low temperatures where the oils are solid or nearly so, but the used oils show a definite maximum at somewhat higher temperatures where the oils become fluid again. The unused and supposedly nonpolar oil shows a very slight decrease in dielectric constant as the temperature is lowered below $-50°C$. This

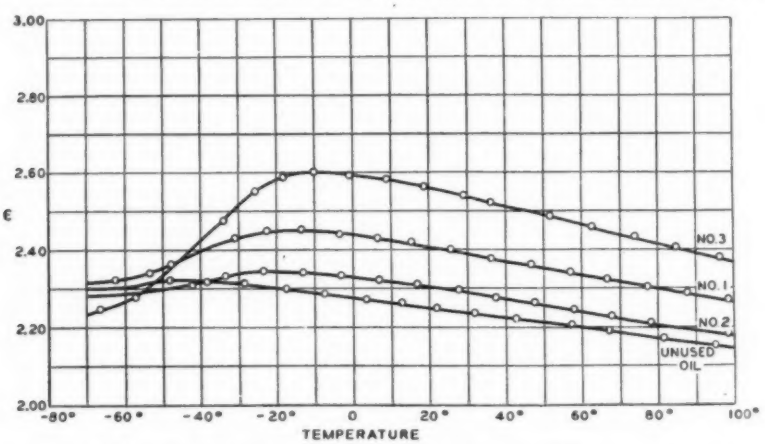



Fig. 6. Variation of dielectric constant of transformer oils with temperature.

may be due to the contraction which occurs when the oil soldifies, with the result that the space between the condenser plates is not completely filled.

The variation of power factor with temperature as shown in Figs. 7 and 8 indicates that it is very largely caused by dipole energy loss as the molecules rotate with the field. The relaxation time, that is, the time in which the or-

derly arrangement of the molecules caused by an electric field will be reduced to $1/e$ of its initial value is defined[4] by the equation $t = 8\pi nr^3/2kT$ where $k$ is the Boltzmann constant, $T$ the absolute temperature, $r$ the radius of the molecule and $n$ the inner friction constant of the liquid. By using the ordinary viscosity for the inner friction constant, it is possible to calculate a characteristic frequency at any temperature for a liquid of known molecular diameter.

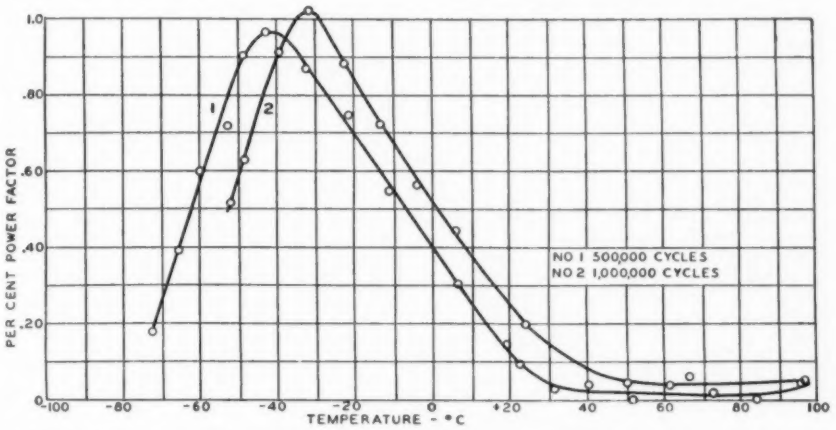



Fig. 7. Variation of power factor of oxidized transformer oil with temperature and frequency.

At frequencies either very much above the characteristic or very much below it, the dipole contribution to the power factor is obviously quite small. In the present case, instead of measuring the power factor at constant temperature over a wide range of frequencies, the more convenient method of varying the temperature was used. Starting at about 100°C, the power factor is practi-

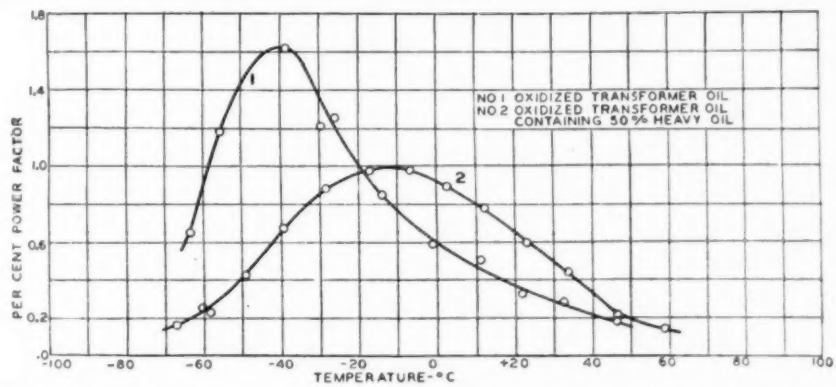



Fig. 8. Variation of power factor of oxidized transformer oils with temperature and viscosity.

cally independent of temperature down to 20°, where a sharp increase begins. As the temperature is lowered still further and the viscosity increases rapidly, the power factor passes through a maximum and drops back to a very small value. When the dipoles are so firmly held that they can no longer fol-

[4] P. Debye. Polar Molecules, p. 84. Chemical Catalog Company, 1929.

low the field, their contribution to the power factor of course becomes negligible. The higher the frequency, the higher the temperatue at which the maximum should occur, and we find the maximum of the 1,000,000 cycle curve about 10° higher than that of the 500,000 cycle curve.

The most interesting feature of Fig. 8 is that in curve No. 1 the maximum occurs at −40°C, about 30° below the freezing temperature of the oil. This shows that although the oil is quite solid, the molecules are still free to rotate with the field, and hence that the mechanical viscosity and the inner friction which opposes the turning motion of the molecules are not always the same. The peak in curve No. 2 is very much broadened and it is shifted from −40° to −10°. The power factor of this more viscous oil is less, due to the dilution with a non-polar oil, and the maximum is shifted to a higher temperature because of the increase in viscosity. Johnstone and Williams[5] have shown that molecular radii of the right order of magnitude are obtained when the experimental values of viscosity and characteristic frequency are substituted in Debye's formula for the relaxation time. However, the fact that the maximum in power factor (and in dielectric constant, although the data are omitted for brevity) occurs many degrees below the solidification temperature of certain oils shows that the inner friction constant or molecular viscosity is not always even approximately equal to the mechanical viscosity.

The measuring of the polarity of transformer oils in this fashion should serve as a useful indication of the extent of decomposition in a given used oil and as a convenient test by which to predict the behavior of oils in transformers. It gives no indication of the tendency of an oil to form sludge except insofar as the degree of oxidation is an indication of the tendency to sludge. The presence of suspended sludge in the oil does not appreciably affect its dielectric constant. However, the method does offer a quick and comparatively easy means of differentiating between several oils as to their tendency to decompose.

The author is pleased to acknowledge his indebtedness to Mr. J. G. Ford for the oils studied and for the data in Fig. 5, and to Dr. C. F. Hill for his helpful advice and criticism throughout the course of the work.

[5] Johnstone and Williams, Phys. Rev. **34,** 1483 (1929).

# Potential Distributions in Large Cylindrical Disks with Partially Penetrating Electrodes

By Morris Muskat
*Gulf Research Laboratory*

(Received March 3, 1932)

Because of the equivalence of the pressure in a liquid bearing sand to a velocity potential, an analysis has been carried through of the potential theory of the electrical analogue of an artesian well partially penetrating a water bearing sand.

**Part I.** In Part I the method of analysis is illustrated by the solution of the problem of a hemispherical electrode imbedded in one of the faces of the disk. Two types of formulae are developed suitable for study of the potential distribution at small and large distances from the electrode respectively. The resistance of this system as a function of the disk thickness is computed and plotted.

**Part II.** In Part II is treated the more general problem of the potential distribution for a partially penetrating electrode, with a uniform flux density along the electrode surface. Separate formulae for the potential function are again derived that are convenient for discussion at small and large separations from the electrode. It is shown that they reduce in the limit of small penetrations to those of Part I and in the limit of complete penetration to well-known radial flow formulae. Expressions are also given for the stream functions corresponding to the derived potential functions. Finally it is shown that when the flux density along the electrode is assumed to be constant the potential over the electrode is far from uniform, falling sharply near the electrode extremity to a value slightly more than half of that at the top of the electrode.

**Part III.** The problem of physical interest in which the electrode is at uniform potential is attacked in Part III. After discussing the case of a flux density over the electrode surface increasing with distance from the top the direct method of superposing discontinuous flux density elements is finally used to obtain electrode surface potentials uniform to within 2 percent of the average surface value. The equivalence of this method to the direct solution of the integral equation for the problem is pointed out. The resistances of the disk for various electrode penetrations and disc thicknesses are then computed and plotted. From these a semi-empirical approximation is derived permitting the resistances for other cases to be derived from a single expression without the necessity of first adjusting the electrode flux distribution so that its potential be uniform.

ALTHOUGH the problem of the potential distributions in cylindrical disks with partially penetrating electrodes appears to have no practical interest electrically, it is of considerable importance when applied to the question of flow of liquids through sands into artesian wells. This significance is due to the fact that the fluid pressure in a sand through which a liquid is flowing, being equivalent to a velocity potential, obeys Laplace's equation. To find the pressure distribution in a sand one has therefore to solve the corresponding potential problem for the same geometrical system.

This analogy may be formulated more precisely in the following way. For a sand bearing a liquid in motion, the velocity potential, $\Phi$, is related to the pressure $p$, at the point $(r, \Theta, z)$ by the equation

$$\Phi = k(+ p - \gamma g z), \tag{1}$$

so that

$$\bar{v} = -\nabla\Phi;\ v_r = -\frac{\partial\Phi}{\partial r},\quad v_\theta = -\frac{1}{r}\frac{\partial\Phi}{\partial\theta},\quad v_z = -\frac{\partial\Phi}{\partial z}, \tag{1a}$$

$$\nabla^2\Phi = 0. \tag{1b}$$

In these equations, $k$ is the permeability constant of the sand for the fluid in question—of density $\gamma-$, and is defined best by Eqs. (1) and (1a) themselves; $g$ is the acceleration of gravity and $\bar{v}$ is the "effective" fluid velocity vector, interpreted as the volumetric fluid flux across a unit section of sand, its components in cylindrical coordinates being $v_r$, $v_\theta$, and $v_z$, $v_\theta$ in the present discussion, always vanishing by symmetry.

Eqs. (1), (1a), and (1b) are really nothing more than the analytical statement of D'Arcy's law established experimentally as an empirical substitute for the Stokes-Navier equations, whose application to the flow in a sand would be a hopelessly difficult and complicated one. Eqs. (1a) further state that $v_r$ and $v_z$ are positive when directed towards positive $r$ and $z$, the latter being taken as downwards.

It is clear that electrically the sand corresponds to a resisting medium of specific conductance $k$, voltage distribution $p$, and current flux distribution $\bar{v}$. Although, from the electrical point of view there is no analogue for the gravity term $\gamma gz$, still, since the problem is sufficiently defined by the potential function $\Phi$, there is no need to specify—except in the final interpretation—whether it may be split up as in Eq. (1) or not. The electrical analogy also implies that a well penetrating a fluid bearing sand corresponds to a conducting electrode and the impermeability of the sand faces corresponds to electrical insulation. Since the present discussion will be essentially a potential theory analysis of the problem in question, the electrical phraseology will be used throughout.

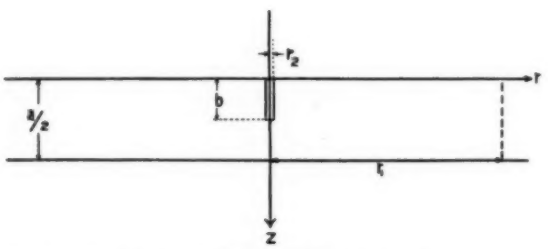



Fig. 1. Diagrammatic representation of disk of radius $r_1$ and thickness $a/2$, penetrated by an electrode of radius $r_2$ and length $b$.

## Statement of the Problem

For simplicity the analysis will be confined to the case of cylindrical disks faced by insulating parallel planes, i.e., such that over them the normal gradient of the potential vanishes. At the center of the disk is inserted a conducting electrode of small radius $r_2$, and length $b$. Corresponding to the case of practical interest when applied to artesian wells, it will be supposed that

the outer radius of the disk, $r_1$, is very large as compared to the electrode radius, $r_2$, and of the same order as or larger than the thickness of the disk, $a/2$ (cf. Fig. 1).

Analytically stated, the problem is: To find a solution of the equation

$$\nabla^2\Phi = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = 0, \tag{1b}$$

such that:

$$\frac{\partial\Phi}{\partial z} = 0,\ z = 0,\ a/2; \tag{1c}$$

$$\left.\begin{aligned}\frac{\partial\Phi}{\partial r} &= \text{const.}\\ \Phi &= \text{const.}\end{aligned}\right\} r = r_2,\ z < b; \tag{1d}$$

or:

$$\Phi = \text{const};\ r = r_1. \tag{1e}$$

Condition (1e) will not be strictly adhered to. Since, as was stated, $r_1$ will be taken as much larger than $r_2$ it follows from elementary considerations that necessarily $\Phi$ will be very approximately constant at $r = r_1$. For all cases of practical interest the slight variation of $\Phi$ with $z$ at $r = r_1$ will be of no importance. Hence in the following analysis no attempt will be made to force $\Phi$ to have a strictly constant value at $r = r_1$ but rather, the $\Phi$ that satisfies the other conditions will be taken as satisfying with sufficient accuracy the last condition also.

It should be mentioned, too, that in the preliminary development of the analysis, as is given in Part II, the first of conditions (1d) will be satisfied. That is, a solution will be developed and discussed, for which the *flux density* along the electrode is uniform; from this, a solution will be developed in Part III such as will satisfy the second of Eqs. (1d) to a close approximation. It is true that only the latter solution is of physical interest, but the first case will be solved essentially as an analytical preliminary problem.

From the symmetry of the problem it is clear that one should use cylindrical coordinates. It will be convenient to place the origin at the center of the upper face and to take $z$ positive as directed downward, as already mentioned. It will also be convenient to introduce from the outset the dimensionless variables:

$$\frac{r}{a} = \rho;\ \frac{z}{a} = w;\ \frac{b}{a} = x;\ \frac{r_2}{a} = \rho_2;\ \frac{r_1}{a} = \rho_1. \tag{2}$$

Before entering on the analysis it should be mentioned that the present problem has recently been investigated by A. F. Samsioe.[1] Since he was not interested in details of the potential distributions, he derived formulae which are not suitable for studying this phase of the problem. Thus, for example, at $r = 0$, his expressions appear to diverge even if $z > b$. Further, for large values of $r$ his equations are rather awkward to handle. Finally, Samsioe has not attempted to satisfy in detail the requirement that the electrode be at a

uniform potential, but treated only the case where the flux density along the electrode is uniform.

Because it has some features not characteristic of the general penetrating electrode and because it turns out to be the most important case for an application made in another paper, it will be useful to discuss first in detail the special problem where the electrode does not penetrate the disk but is embedded in its upper face. This problem will therefore be solved first.

## I. The Nonpenetrating Electrode

### A. Potential function for small values of $\rho$

By a nonpenetrating electrode is simply meant a semispherical cap embedded in the upper face of the disk. For purposes of the analysis, however, it will be convenient to replace the electrode by a small spherical equipotential surface enveloping a mathematical source of strength $q$, placed at the origin. The fundamental solution of the potential equation will therefore have the form:

$$\Phi = \frac{q}{(r^2 + z^2)^{1/2}} = \frac{q}{a(\rho^2 + w^2)^{1/2}} . \tag{3}$$

To obtain a solution satisfying Eq. (1c) the well-known method of images, also used by Samsioe, will be employed. Thus, it will be readily verified that if images of the source at the origin be placed along the axis at the points $(o, \pm na)$ the resulting potential distribution will give zero flux across the planes $z=0, a/2$. With this arrangement the problem is formally solved, and there remains only the development of suitable formulae to represent the solution.

Thus, from Eq. (3) it is seen that the complete solution has the form:

$$\Phi = \frac{q}{a} \sum_0^\infty n \frac{1}{(\rho^2 + (n + w)^2)^{1/2}} + \frac{q}{a} \sum_1^\infty n \frac{1}{(\rho^2 + (n - w)^2)^{1/2}} + \text{const.} \tag{4}$$

Looking at Eq. (4), without reference to Eqs. (1c) and (1d), it will be noted that it represents the potential due to an infinite regular array of equal point sources or charges, lying on the $z$-axis, and spaced by the distance $a$. From this point of view Eq. (4) has been discussed by Madelung[2] as a preliminary to his analysis of the potential fields in space lattices of heteropolar crystals. Although Madelung does not concern himself, in the case of the line array, with details of the potential distribution, it will be convenient to make use of some of his transformations.

As is to be expected of the potential due to an infinite set of like images, Eq. (4), as it stands, is everywhere divergent and appears therefore to have no meaning. However, this difficulty is only apparent and of no physical consequence since one may subtract an arbitrarily large constant from the summations and in the limit obtain a convergent representation of the variable field at all points of interest. This will appear more clearly presently.

[1] Samsioe, Zeits. f. ang. Math. u. Mech. **11**, 124 (1931).

[2] Madelung, Phys. Zeits. **19**, 524 (1918).

Thus, neglecting for the moment the question of convergence, the series in Eq. (4) can be expanded for: $\rho<1-w\geq 1/2$, i.e., for small values of $r$, as:

$$\left.\begin{aligned}\Phi = {} & \frac{q}{a(\rho^2+w^2)^{1/2}} + \frac{q}{a}\sum_1^\infty{}^n \frac{1}{n+w}\left\{1-\frac{1}{2}\frac{\rho^2}{(n+w)^2}\right. \\ & \left.+\frac{3}{8}\frac{\rho^4}{(n+w)^4}\cdots\right\} + \frac{q}{a}\sum_1^\infty{}^n \frac{1}{n-w}\left\{1-\frac{1}{2}\frac{\rho^2}{(n-w)^2}\right. \\ & \left.+\frac{3}{8}\frac{\rho^4}{(n-w)^4}-\cdots\right\} + \text{const.}\end{aligned}\right\} \tag{5}$$

Now, from the theory of gamma-functions one has the relations:

$$\left.\begin{aligned}\psi(y) = \frac{\Gamma'(y)}{\Gamma(y)} = {} & -\delta - \frac{1}{y} + \lim_{n\to\infty}\sum_1^n{}^m\left(\frac{1}{m}-\frac{1}{y+m}\right) \\ \text{where}\qquad \delta = {} & \lim_{n\to\infty}\left(\sum_1^n{}^m \frac{1}{m} - \log n\right) = 0.5772.\end{aligned}\right\} \tag{6}$$

Hence:

$$\left.\begin{aligned}\sum_1^\infty{}^n \frac{1}{n+w} &= -\psi(w) - \frac{1}{w} + \lim_{n\to\infty}\log n, \\ \sum_1^\infty{}^n \frac{1}{n-w} &= -\psi(1-w) + \lim_{n\to\infty}\log n.\end{aligned}\right\} \tag{6a}$$

Further, by the definition of the $\zeta$-functions, one has:

$$\left.\begin{aligned}\sum_1^\infty{}^n \frac{1}{(n+w)^s} &= \zeta(s,w) - \frac{1}{w^s}, \\ \sum_1^\infty{}^n \frac{1}{(n-w)^s} &= \zeta(s,1-w).\end{aligned}\right\} \tag{6b}$$

Eq. (5) can therefore be rewritten in the form:

$$\left.\begin{aligned}\Phi = \frac{q}{a}\Bigg[ & \frac{1}{(\rho^2+w^2)^{1/2}} - \psi(w) - \frac{1}{w} - \psi(1-w) - \frac{1}{2}\rho^2\left\{\zeta(3,w)\right. \\ & \left.-\frac{1}{w^3}+\zeta(3,1-w)\right\} + \frac{3}{8}\rho^4\left\{\zeta(5,w) - \frac{1}{w^5}\right. \\ & \left.+\zeta(5,1-w)\right\} - \cdots + 2\lim_{n\to\infty}\log n\Bigg] + \text{const.}\end{aligned}\right\} \tag{7}$$

From this expression it is seen that the divergence of the series in Eq. (4) may be separated out into the term $2\lim_{n\to\infty}(\log n)$ which can be formally removed by making the arbitrary constant term equal and opposite to it. It will therefore be dropped hereafter although a finite constant term may be added so as to make possible an arbitrary choice of $\Phi$ at some given point.

Since the $\psi$-functions are usually tabulated only for arguments greater than 1 it is useful to transform the functions in Eq. (7) so that they satisfy this condition. This may be done by means of the well-known relations:

$$\psi(y) = -\frac{1}{y} + \psi(1+y);\ \psi(1-y) = \psi(y) + \pi \operatorname{ctn} \pi y. \tag{6c}$$

Applying these to Eq. (7) one gets finally:

$$\left.\begin{aligned}\Phi = \frac{q}{a}\Big[ &-2\psi(1+w) + \frac{1}{w} - \pi \operatorname{ctn} \pi w + \frac{1}{(\rho^2+w^2)^{1/2}} \\ &-\frac{1}{2}\rho^2\left\{\zeta(3,w) - \frac{1}{w^3} + \zeta(3,1-w)\right\} \\ &+\frac{3}{8}\rho^4\left\{\zeta(5,w) - \frac{1}{w^5} + \zeta(5,1-w)\right\} - \cdots\Big].\end{aligned}\right\} \tag{8}$$

It is readily shown that for very small distances from the origin Eq. (8) reduces to:

$$\Phi = \frac{q}{a}\cdot\frac{1}{(\rho^2+w^2)^{1/2}};\ (\rho^2+w^2)^{1/2} \ll 1. \tag{8a}$$

This means physically that very near the origin the potential is spherical and due almost entirely to the single source at the origin, and this is, of course, what is to be required if the electrode is to have a spherical shape.

From Eq. (8a) one can readily obtain the total flux $Q$ leaving the electrode and flowing through the disk. Thus:

$$Q = \int \frac{\partial \Phi}{\partial n}\, d\sigma = +2\pi q, \tag{9}$$

where in the integral, $n$ is the normal at the element $d\sigma$ on any surface enclosing the origin, the integral extending, however, only over that part of the surface lying between the planes $z=0$, and $z=a/2$. Eq. (9) simply verifies our interpretation of $q$ as the source "strength".

**B. Potential function for large values of $\rho$**

For small values of $\rho$ Eq. (8) can be handled numerically without difficulty. For large values of $\rho$, Eq. (8) diverges; but even the corrected series—obtained by separating out terms like $1/(\rho^2+w^2)^{1/2}$ from the series in Eq. (4)—is quite unsuitable for detailed study. In fact, one must abandon the power series in $\rho$. As a substitute it will be convenient to make use of Madelung's rapidly convergent series obtained by a direct solution of the potential equation for a continuously distributed charge along the $z$-axis, and so adjusted that it goes over into the potential due to a point array on properly passing to the limit of a concentrated charge density.

The equivalent of Eq. (8) becomes, by Madelung's representation:

$$\Phi = \frac{q}{a}\left[4\sum_{1}^{\infty} K_0(2\pi n\rho)\cos 2\pi n w + 2\log\frac{2}{\rho}\right] \tag{10}$$

where $K_0$ is the Hankel function of order 0. For $\rho > 0.8$ the first term in the sum suffices for almost all purposes. Again, in Eq. (10) an infinite constant has been dropped as it was in Eq. (8).

It is readily seen from Eq. (10) that $\Phi$ becomes practically independent of $z$ as soon as $r$ becomes of the order of $a$, i.e., the flux is essentially radial except in the region near the electrode no more distant than about twice the thickness of the disk. This will be seen more clearly from a plot of $\Phi$ to be presented later, and it justifies in detail the remarks made on page 331, that solutions satisfying Eqs. (1c) and (1d) will necessarily satisfy Eq. (1e) with sufficient accuracy.

It is of interest to derive here the relation between the total flux $Q$ and the potential difference $\Delta\Phi$ existing between the electrode and some large distance $r_1$ from it. This may be found as follows. As:

$$K_0(y) \leqq \left(\frac{\pi}{2y}\right)^{1/2} e^{-y}$$

for large values of $y$ one may omit the summation in Eq. (10) for large $\rho$; whence it reduces to:

$$\Phi = \frac{2q}{a}\log 2/\rho. \tag{10a}$$

Now it has already been noted that for small distances from the origin $\Phi$ depends only on $(\rho^2+w^2)^{1/2}$. Hence specifying that the electrode radius is given by $(\rho^2+w^2)^{1/2}=\rho_2$, the potential of the electrode may be expressed by:

$$\Phi = \Phi(w=\rho_2,\ \rho=0) = \frac{q}{a}\left[-2\psi(1+\rho_2) + \frac{2}{\rho_2} - \pi\operatorname{ctn}\pi\rho_2\right].$$

The potential drop then becomes:

$$\Delta\Phi = \frac{-q}{a}\left[2\log\frac{2}{\rho_1} + 2\psi(1+\rho_2) - \frac{2}{\rho_2} + \pi\operatorname{ctn}\pi\rho_2\right]. \tag{11}$$

Now $\Delta\Phi/Q$ is clearly the electrical resistance of the disk for unit conductance of the disk material. Denoting this resistance by $\Omega$, it follows from Eqs. (11) and (9) that:

$$\Omega = \frac{\Delta\Phi}{Q} = \frac{1}{2\pi a}\left[\frac{2}{\rho_2} - \pi\operatorname{ctn}\pi\rho_2 - 2\psi(1+\rho_2) - 2\log\frac{2}{\rho_1}\right]. \tag{12}$$

Noting that for $\rho_2 \ll 1$, $\pi\operatorname{ctn}\pi\rho_2 \leqq 1/\rho_2$ Eq. (12) may be approximated by:

$$\Omega = \frac{1}{2\pi r_2} - \frac{1}{\pi a}(\psi(1) + \log 2/\rho_1). \tag{12a}$$

Hence, since the last two terms are evidently smaller than the first, except for small values of $a$, the resistance for the nonpenetrating electrode is practically independent of the thickness of the disk. To show this more clearly the value of $\Omega$ is plotted in Fig. 2 as a function of $a$ for the case where $r_2=0.01$ and $r_1=10$. The range of $a$ is from $1/2$ to 10 so that $\rho_2$ varies from 0.02 to 0.001.

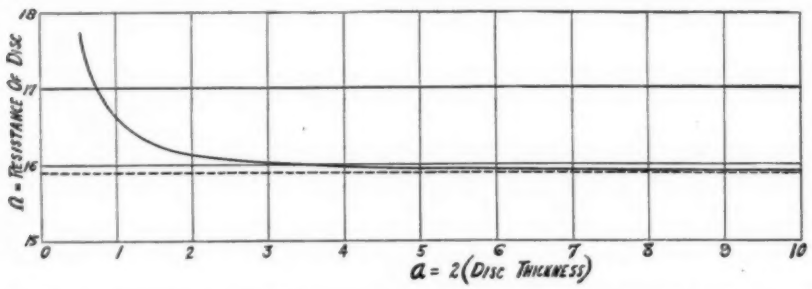



Fig. 2. Resistance of disk with a nonpenetrating electrode as a function of the thickness of the disk. Electrode radius $r_2=0.01$ and disk radius $r_1=10.0$. Horizontal dotted line represents resistance of a semi-spherical shell of inner and outer radii of 0.01 and 10.0 respectively.

One may note, too, that the last terms in (12a) simply represent the deviation of $\Omega$ from that for a perfectly semispherical resisting shell, since the resistance of the latter may be easily shown to be:

$$\Omega = \frac{1}{2\pi}\left(\frac{1}{r_2} - \frac{1}{r_1}\right) \leqq \frac{1}{2\pi r_2}. \tag{12b}$$

The horizontal line in Fig. 2 is the value of $\Omega$ for purely semispherical flow and again with $r_2=0.01$, $r_1=10$. It is seen from this that as long as $\rho_2$

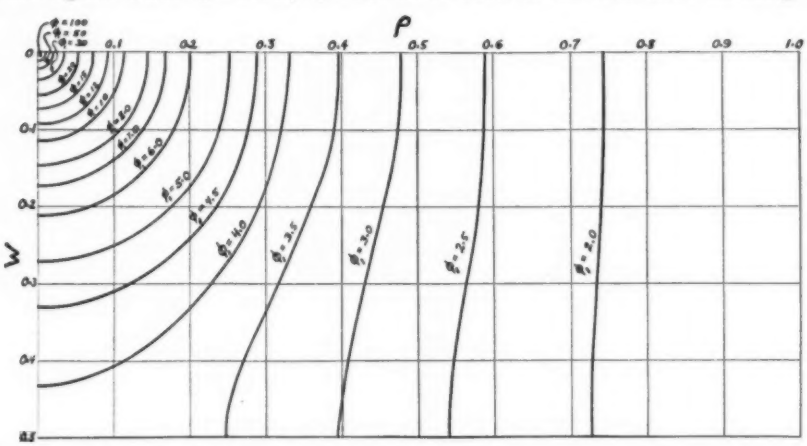



Fig. 3. Equipotential curves in a disk with a nonpenetrating electrode placed at center of its top surface. $\Phi_1$ is the potential function of equation (13). $\rho$ = radial distance /2 (disk thickness). $w$ = (depth in disk) /2 (disk thickness).

$=r_2/a<0.005$ the resistance of the disk differs but slightly from that of the corresponding hemisphere.

To see more clearly the nature of the potential distribution both for small distances from the electrode and distances at which the radial character of the flux begins to predominate, a plot of several equipotential curves in a plane through the $z$-axis is given in Fig. 3. To make the distribution independent of any particular thickness of the disk the latter was eliminated by taking $q/a=1$ in Eq. (8). The particular electrode radius was also eliminated by leaving the absolute potential over the electrode quite arbitrary. Thus the potential function plotted in Fig. 3 is given by

$$\left.\begin{aligned}\Phi_1 = {} & \frac{1}{(\rho^2+w^2)^{1/2}}+\frac{1}{w}-\pi\operatorname{ctn}\pi w-2\psi(1+w)\\ & -\frac{1}{2}\rho^2\left\{\zeta(3,w)-\frac{1}{w^3}+\zeta(3,1-w)\right\}+\frac{3}{8}\rho^4\left\{\zeta(5,w)\right.\\ & \left.-\frac{1}{w^5}+\zeta(5,1-w)\right\}\cdots;\ \rho<\tfrac{1}{2};\\ \Phi_1 = {} & 4\sum_1^\infty{}' K_0(2\pi n\rho)\cos 2\pi nw+2\log 2/\rho:\rho>\tfrac{1}{2}.\end{aligned}\right\}\qquad(13)$$

Although tables of the $\zeta$-functions were available for use in the first value of $\Phi_1$ for $\rho<1/2$, it was found just as convenient to use the second form with three terms in the sum even for $\rho=0.3$.

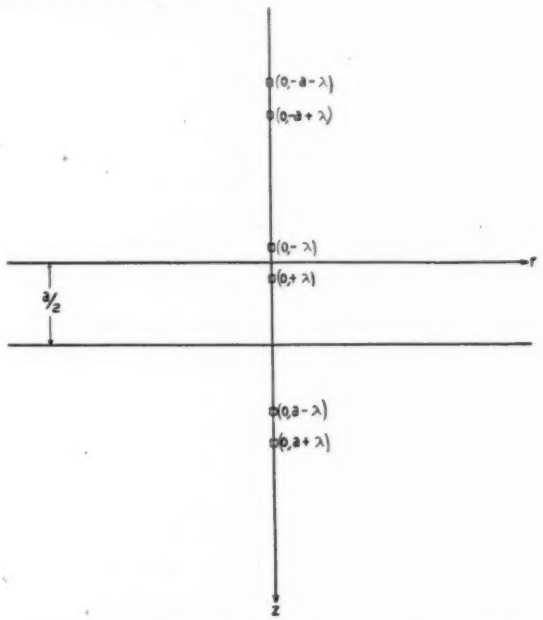



Fig. 4. Diagrammatic scheme of image system used to represent potential functions for partially penetrating electrodes.

It will be noted that, as already indicated before, although near the electrode the equipotential surfaces are spherical, just as about a single point

charge, as soon as one recedes from the electrode to a distance less even than twice the thickness of the disk the equipotential surfaces become closely cylindrical as if the electrode penetrated the disk completely.

For the present purposes the above discussion will suffice for describing the features peculiar to the special problem of the nonpenetrating electrode. As to the other details, it will be better to bring them out in comparison with the cases of penetrating electrodes. These will therefore be analyzed in the following section.

## II. Partially Penetrating Electrodes With Uniform Flux Density

### A. Potential function for small values of $\rho$

To analyze the case of the partially penetrating electrodes, it is convenient to begin with the potential function due to an element of source distribution located along the $z$-axis at a distance $\lambda$ from the origin. The elementary potential function in this case is of the form:

$$\Phi = \frac{qd\lambda}{(r^2 + (z - \lambda)^2)^{1/2}} = \frac{qd\mu}{(\rho^2 + (w - \mu)^2)^{1/2}}, \tag{14}$$

where $qd\lambda$ is the strength of the element of length $d\lambda$, and $\mu$ is the dimensionless variable $\lambda/a$.

Again, to obtain a function $\Phi$ which will satisfy Eq. (1c) it is necessary to add to the original source at $(0, \lambda)$ an appropriate set of images. These, it will be readily verified, should be placed at the points $(0, \pm na \pm \lambda)$. One therefore has, for the differential potential function due to the element at $\lambda$ and its images, the value:

$$\left.\begin{aligned} d\Phi = qd\mu\Bigg[&\frac{1}{(\rho^2 + (w - \mu)^2)^{1/2}} + \frac{1}{(\rho^2 + (w + \mu)^2)^{1/2}} \\ &+ \sum_1^\infty{}^* \left\{ \frac{1}{(\rho^2 + (n + w - \mu)^2)^{1/2}} + \frac{1}{(\rho^2 + (n + w + \mu)^2)^{1/2}} \right. \\ &\left. + \frac{1}{(\rho^2 + (n - w - \mu)^2)^{1/2}} + \frac{1}{(\rho^2 + (n - w + \mu)^2)^{1/2}} \right\} \Bigg]. \end{aligned} \right\} \tag{15}$$

* To be exact the first term in Eq. (15) should be changed to $1/(\rho^2+(\mu-w)^2)^{1/2}$ for $w<\mu$. However, this is of no significance except when $\rho=0$. In the following, therefore, Eq. (15) will be used as it is, since a detailed check will show that all the resulting formulae are rigorously correct. It should be pointed out, too, that the author does not pretend that the formal solution in terms of images, as expressed by Eq. (15) is new or original for as early as 1873 Maxwell set up this same system for a point electrode in a disk and even for the more general case where the disk is surrounded by other conducting media. Further, geophysicists interested in electrical prospecting have recently worked on this more general problem of Maxwell. However, no detailed development and discussion of Eq. (15) or of its equivalent for large values of $r$ as presented here has been found in the literature, except in the references specifically cited, nor has a treatment of the partially penetrating electrode been given by anyone except by Samsioe, as already mentioned.

Supposing now that the source at $\lambda$ is continued with constant strength along the $z$-axis from $y=0$ to $y=b$, i.e., from $\mu=0$ to $\mu=x$, the potential due to the continuous and uniform distribution will be given by:

$$\left.\begin{aligned}\Phi &= \int_0^b d\Phi \\ &= q\Big[\sum_1^\infty{}^n \log \frac{n-w+x+(\rho^2+(n-w+x)^2)^{1/2}}{n-w-x+(\rho^2+(n-w-x)^2)^{1/2}} \\ &\quad + \sum_0^\infty{}^n \log \frac{n+w+x+(\rho^2+(n+w+x)^2)^{1/2}}{n+w-x+(\rho^2+(n+w-x)^2)^{1/2}}\Big].\end{aligned}\right\} \tag{16}$$

This is essentially the procedure used by Samsioe. However, his counting of the images led him to log terms whose arguments vanish for $\rho=0$, although one could easily transform his expression so as to be equivalent to those in (16), before passing to the limit $\rho=0$. In Eq. (16), $\Phi$ has a singularity at $\rho=0$ only for $w<x$, for then the term $\log\{w-x+(\rho^2+(w-x)^2)^{1/2}\}$ and hence $\Phi$ becomes infinite as $\log\rho$, as indeed it should.

Nevertheless, Eq. (16) is not particularly convenient to handle unless one goes through lengthy expansions of the individual logarithmic terms. It is therefore better to expand the series in Eq. (15) directly, as was done in the case of the nonpenetrating electrode and then integrate. By this method, then, one obtains first: for $\rho<1-x-w$, to terms in $\rho^6$,

$$\left.\begin{aligned}d\Phi = qd\mu\Big[&\frac{1}{(\rho^2+(w-\mu)^2)^{1/2}}+\frac{1}{(\rho^2+(w+\mu)^2)^{1/2}} \\ &+\sum_1^\infty{}^n\Big\{\frac{1}{n-w-\mu}\Big(1-\frac{1}{2}\frac{\rho^2}{(n-w-\mu)^2}+\frac{3}{8}\frac{\rho^4}{(n-w-\mu)^4}\Big) \\ &+\frac{1}{n-w+\mu}\Big(1-\frac{1}{2}\frac{\rho^2}{(n-w+\mu)^2}+\frac{3}{8}\frac{\rho^4}{(n-w+\mu)^4}\Big) \\ &+\frac{1}{n+w+\mu}\Big(1-\frac{1}{2}\frac{\rho^2}{(n+w+\mu)^2}+\frac{3}{8}\frac{\rho^4}{(n+w+\mu)^2}\Big) \\ &+\frac{1}{n+w-\mu}\Big(1-\frac{1}{2}\frac{\rho^2}{(n+w-\mu)^2}+\frac{3}{8}\frac{\rho^4}{(n+w-\mu)^4}\Big)\Big\}\Big].\end{aligned}\right\} \tag{17}$$

The restriction $\rho<1-x-w$ can be advanced to $\rho<1-w$ or $\rho<1-x$ by removing in the first summation in Eq. (17) the terms for $n=1$ and separating them out into the single term $1/(\rho^2+(1-w-\mu)^2)^{1/2}$, but since in the application and discussion to be made later the series expansion for $\Phi$ will be used only for $\rho\ll 1$, it will be somewhat easier to use Eq. (17) as it stands.

Now by the aid of Eqs. (6a) and (6b) for the $\psi$ and $\zeta$-functions, Eq. (17) may be written as:

$$
\begin{aligned}
d\Phi = q d\mu \Bigg[ & \frac{1}{(\rho^2 + (w-\mu)^2)^{1/2}} + \frac{1}{(\rho^2 + (w+\mu)^2)^{1/2}} \\
& - \psi(1 - w - \mu) - \psi(1 - w + \mu) - \psi(w + \mu) - \psi(w - \mu) \\
& - \frac{1}{w+\mu} - \frac{1}{w-\mu} - \frac{1}{2}\rho^2 \Bigg\{ \zeta(3, 1 - w - \mu) - \frac{1}{(w+\mu)^3} \\
& + \zeta(3, w + \mu) + \zeta(3, 1 - w + \mu) - \frac{1}{(w-\mu)^3} \\
& + \zeta(3, w - \mu) \Bigg\} + \frac{3}{8}\rho^4 \Bigg\{ \zeta(5, 1 - w - \mu) - \frac{1}{(w+\mu)^5} \\
& + \zeta(5, w + \mu) + \zeta(5, 1 - w + \mu) - \frac{1}{(w-\mu)^5} \\
& + \zeta(5, w - \mu) \Bigg\} - \cdots + 4 \lim_{n \to \infty} \log n \Bigg].
\end{aligned}
\tag{17a}
$$

Again, the divergence of the series in Eq. (15) appears to be in the term $4 \lim_{n\to\infty} \log n$, and for the same reasons as already discussed this term will be dropped in the following.

As the potential $d\Phi$ will have to be integrated with respect to $\mu$ the various terms must be expressed in differential form. This may be done on noting the following relations:

$$
\begin{aligned}
\psi(1 - w - \mu) &= \psi(w + \mu) + \pi \operatorname{ctn} \pi(w + \mu), \\
\psi(w - \mu) &= \psi(1 - w + \mu) - \pi \operatorname{ctn} \pi(w - \mu), \\
\psi(m + \mu) &= \frac{d}{d\mu} \log \Gamma(m + \mu), \\
\zeta(s, \mu) &= \frac{1}{1 - s} \frac{\partial}{\partial \mu} \zeta(s - 1, \mu).
\end{aligned}
\tag{18}
$$

Integrating now Eq. (17a) and making use of the well-known relation:

$$
\Gamma(w)\Gamma(1 - w) = \frac{\pi}{\sin \pi w}, \tag{18a}
$$

one obtains finally:

$$
\begin{aligned}
\Phi &= \int_0^x \frac{\partial \Phi}{\partial \mu} d\mu \\
&= q \Bigg[ 2 \log \pi - 2 \log \Gamma(w + x)\Gamma(1 - w + x) \\
&\quad - \log \sin \pi(w + x) \sin \pi(w - x \ + \log \frac{w - x}{w + x}
\end{aligned}
$$

$$+\log\frac{w+x+(\rho^2+(w+x)^2)^{1/2}}{w-x+(\rho^2+(w-x)^2)^{1/2}}-\frac{1}{4}\rho^2\Big\{\zeta(2,1-w-x)-\zeta(2,1-w+x)+\zeta(2,w-x)-\zeta(2,w+x)+\frac{1}{(w+x)^2}-\frac{1}{(w-x)^2}\Big\}+\frac{3}{32}\rho^4\Big\{\zeta(4,1-w-x)-\zeta(4,1-w+x)+\zeta(4,w-x)-\zeta(4,w+x)+\frac{1}{(w+x)^4}-\frac{1}{(w-x)^4}\Big\}-\cdots\Big] \quad (19)$$

This is the final form for the potential; it is useful numerically for small values of $\rho$ and valid only for *constant flux density, $q$, along the electrode.* It will be discussed in detail in sections **C** and **D**. For large values of $\rho$ another expression will be derived presently.

### B. Potential functions for large values of $\rho$

Since Eq. (19) is essentially a power series in $\rho$, it is clear that it is not suitable for studying $\Phi$ for large values of $\rho$, i.e., at large distances from the electrode. A suitable expression can be obtained, however, by the following extension of Madelung's expression for the potential due to a point lattice of equal charges, as is given in Eq. (10). Noting that the system of images as indicated in Fig. 4 is equivalent to a superposition of two such lattices, it is readily verified, on comparing with Eq. (10), that the potential due to the source at $\lambda$ and its set of images, is given by:

$$d\Phi = qd\mu\Big[4\sum_1^\infty{}_n K_0(2\pi n\rho)\{\cos 2\pi n(w-\mu)+\cos 2\pi n(w+\mu)\}+4\log\frac{2}{\rho}\Big]. \quad (20^*)$$

* The extended character of this whole analysis may seem artificial when it is observed that a direct solution of Laplace's equation for the present problem gives the simple results for a flux element $qd\mu$ at $\mu$:

$w \leqq \mu$

$$\Phi = 2q(\mu)d\mu\int_0^\infty \frac{J_0(\rho\lambda)\cosh\lambda w\cosh\lambda(\tfrac{1}{2}-\mu)}{\sinh\frac{\lambda}{2}}d\lambda$$

$w \geqq \mu$

$$\Phi = 2q(\mu)d\mu\int_0^\infty \frac{J_0(\rho\lambda)\cosh\lambda\mu\cosh\lambda(\tfrac{1}{2}-w)}{\sinh\frac{\lambda}{2}}d\lambda.$$

However, a little inspection will show that a detailed discussion of these integrals would lead to an analysis and results quite equivalent to the lengthy developments used here. In fact it is not difficult to show by contour integration that the above integrals have an expansion in terms of the Hankel functions identical with Eq. (20). The use of these integrals as a starting point for a detailed discussion of the problem would therefore not shorten the work over the somewhat more direct and more familiar method of images used in the text of this paper, for

Integrating now from $\mu=0$ to $\mu=x$, with the assumption that $q$ is constant, one readily finds as the equivalent of Eq. (19):

$$\Phi = 4q\left[\frac{1}{\pi}\sum_{1}^{\infty}{}'\frac{1}{n}K_0(2\pi n\rho)\cos 2\pi nw \sin 2\pi nx + x\log\frac{2}{\rho}\right]. \quad (21)$$

Again, as the Hankel function $K_0$ decreases exponentially for large arguments, Eq. (21) is very easy to handle numerically even for relatively small values of $\rho$. Eq. (21) also permits an easy calculation of the total flux through the disk, which is given by

$$Q = \int \frac{\partial\Phi}{\partial n}\,d\sigma = -2\pi a\int_0^{1/2}\rho\frac{\partial\Phi}{\partial\rho}\,dw, \quad (22)$$

where $n$ is the normal to the element $d\sigma$ of the surface enveloping the electrode. The second integral implies that the enveloping surface is a cylinder coaxial with the electrode. Applying then Eq. (21) to (22) it is seen that the trigonometric terms give no contribution and the result is obtained that:

$$Q = 4\pi aqx = 4\pi qb, \quad (22a)$$

which again is really nothing more than a confirmation of our physical interpretation of the coefficient $q$.

Eq. (21) shows further that for appreciable distances from the electrode the potential distribution, just as in the case of the nonpenetrating electrode, takes on a radial character and loses its dependence on $z$. And as is to be expected, this approach to a radial flow character is more rapid in the case of the penetrating electrode. This is seen from Eq. (21) in the fact that each term depending on $w$, and hence on $z$, has a factor $\sin 2n\pi x/n<1$ in addition to the Hankel function which alone occurs in the corresponding Eq. (10) for a nonpenetrating electrode.

### C. Limiting cases of zero and complete penetration

Before considering the details of the potential distribution it will be of interest to examine two limiting cases with respect to the electrode penetration. Thus, first, it will be shown that it reduces to the well-known radial flow potential, *independent of* $w$, when the electrode completely penetrates the disk, i.e., when $b=a/2$, $x=0.5$.

This may be done as follows: Noting that:

$$\left.\begin{aligned} &\zeta(s, y) - \zeta(s, y+1) = \frac{1}{y^s} \\ &\text{and:} \\ &\log\{w + (\rho^2 + w^2)^{1/2}\} = \log 2w + \frac{1}{4}\frac{\rho^2}{w^2} - \frac{3}{32}\frac{\rho^4}{w^4} + \cdots;\ w > \rho > 0, \end{aligned}\right\} \quad (18b)$$

even the image expansion of Eq. (15) can be also obtained directly from these integrals on expressing csch $(\lambda/2)$ as a series of exponentials and then integrating term by term. The elegance of the integral representation of $\Phi$ is therefore only apparent, at least as far as a detailed discussion of $\Phi$ is concerned.

one readily finds that the series in $\rho$ in Eq. (19), involving the $\zeta$-functions, reduces to the value:

$$2 \log 2 + \log (w + \tfrac{1}{2})(\tfrac{1}{2} - w) - \log \left\{ w + \tfrac{1}{2} + (\rho^2 + (w + \tfrac{1}{2})^2)^{1/2} \right\} \left\{ \tfrac{1}{2} - w + (\rho^2 + (\tfrac{1}{2} - w)^2)^{1/2} \right\}.$$

With regard to the remainder of Eq. (19) one makes use of Eq. (18a) and then after a slight manipulation it reduces to:

$$\Phi = 2q \log \frac{2}{\rho}; \; x = \tfrac{1}{2}, \tag{23}$$

which is, in fact, the exact value that one should have for pure radial flow from a uniform continuous source along the axis of the disk.

The other limiting case of interest is that where the penetration of the electrode is very small and approaches the case of the nonpenetrating electrode. It will therefore be shown now that for $x$ very small, the general Eq. (19) does reduce to the expression already found for the nonpenetrating electrode. To do this, it is useful to note first the following relations, valid for small values of $x$:

$$\log \Gamma(y + x) = \log \Gamma(y) + x\psi(y);$$

$$\log \sin \pi(y \pm x) = \log \sin \pi y \pm x\pi \operatorname{ctn} \pi y;$$

$$\log \frac{y + x + (\rho^2 + (y + x)^2)^{1/2}}{y - x + (\rho^2 + (y - x)^2)^{1/2}} = \log \frac{y + x}{y - x} - \frac{2x}{y} + \frac{2x}{(\rho^2 + y^2)^{1/2}};$$

$$\frac{1}{(y + x)^s} - \frac{1}{(y - x)^s} = - \frac{2xs}{y^{s+1}};$$

$$\zeta(s, y \pm x) = \zeta(s, y) \mp sx\zeta(s + 1, y).$$

Putting these equations into Eq. (19) and again making use of Eq. (6) one obtains finally:

$$\left.\begin{aligned} \Phi = 2qx \Bigg[ & \frac{1}{(\rho^2 + w^2)^{1/2}} - \psi(w) - \frac{1}{w} - \psi(1 - w) \\ & - \frac{1}{2} \rho^2 \left\{ \zeta(3, w) - \frac{1}{w^3} + \zeta(3, 1 - w) \right\} + \frac{3}{8} \rho^4 \left\{ \zeta(5, w) \right. \\ & \left. - \frac{1}{w^5} + \zeta(5, 1 - w) \right\} - \cdots \Bigg]. \end{aligned}\right\} \tag{24}$$

The expression in the bracket is clearly identical with the corresponding expressions for the case of the nonpenetrating electrode as is given in Eqs. (7) and (8). Just as in this latter case the factor $q/a$ represented the quantity $Q/2\pi a$, where $Q$ is the total flux from the electrode, so, here the factor $2qx$ is also equivalent to $Q/2\pi a$, as is seen from Eq. (22a).

The limiting values of the potential function are still more easily derived from the expression in Eq. (21) developed for use for large values of $\rho$. Thus

for the case of complete penetration one has only to set $x=1/2$ in Eq. (21) to obtain at once:

$$\Phi = 2q \log 2/\rho,$$

as was found from the series expression for $\Phi$ (cf. Eq. (23).)

On the other hand for the nonpenetrating case, i.e., $x$ small, one need only note that

$$\sin 2n\pi x \leqq 2n\pi x : x \text{ small}.$$

This may be assumed without appreciable error, even for $n$ large since then $K_0(2\pi n\rho)$ is anyway so small that the contribution of the whole term is negligible. With the above substitution Eq. (21) becomes:

$$\Phi = 2qx\left[4\sum_1^\infty {}_n K_0(2n\pi\rho)\cos 2n\pi w + 2\log\frac{2}{\rho}\right],$$

which again is identical with Eq. (10) taking into account Eq. (22a).

Hence it is seen that in the limits of very small penetrations the general case of the partially penetrating electrode goes over continuously into that for the nonpenetrating case, and in the other limit of complete penetration, the limit of pure radial flow is also approached continuously.

**D. Stream line distributions**

It is often convenient, in potential problems, to study the stream lines directly rather than by inference from the equi-potential curves. It is therefore of some interest to have available the explicit expressions for the stream functions corresponding to the potential function formulae, although they will not be discussed further in the present paper.

In terms of the present notation the velocity components are related to the stream function $\psi(\rho, w)$ by the equations:

$$\left.\begin{aligned} v_r &= \frac{1}{a^2\rho}\frac{\partial\psi}{\partial w} = -\frac{1}{a}\frac{\partial\Phi}{\partial\rho} \\ v_z &= -\frac{1}{a^2\rho}\frac{\partial\psi}{\partial\rho} = -\frac{1}{a}\frac{\partial\Phi}{\partial w} \end{aligned}\right\} \tag{25}$$

so that, in terms of the potential function $\Phi$:

$$\psi(\rho, w) = -a\rho\int\frac{\partial\Phi}{\partial\rho}dw = a\int\rho\frac{\partial\Phi}{\partial w}d\rho. \tag{26}$$

The results of evaluating the integrals of Eq. (26) by means of previously developed formulae for $\Phi$ are: Flux element $q\,d\mu$ at $(o, \mu)$; potential functions: Eqs. (17a), (20):

$\rho \ll 1$

$$\psi(\rho, w) = aqd\mu\left[\frac{w-\mu}{(\rho^2+(w-\mu)^2)^{1/2}} + \frac{w+\mu}{(\rho^2+(w+\mu)^2)^{1/2}}\right] + 0(\rho^2) \tag{27}$$

$\rho \geqq 1$

$$\psi(\rho, w) = 4aqd\mu\left[w + 2\rho \sum_1^\infty{}_n K_1(2n\pi\rho) \sin 2n\pi w \cos 2n\pi\mu\right]. \tag{27a}$$

Flux segment: $q(\mu) = 1$ $0 \leqq \mu \leqq x$; potential functions: Eqs. (19), (21).

$\rho \ll 1$

$$\psi(x, \rho, w) = a[(\rho^2 + (w + x)^2)^{1/2} - (\rho^2 + (w - x)^2)^{1/2}] + 0(\rho^2) \tag{28}$$

$\rho \geq 1$

$$\psi(x, \rho, w) = 4a\left[xw + \frac{\rho}{\pi} \sum_1^\infty{}_n \frac{K_1(2n\pi\rho)}{n} \sin 2n\pi w \sin 2n\pi x\right] \tag{28a}$$

Of course, as in the case of the potential functions, the trigonometric series for $\psi$, as in Eqs. (27a) and (28a), are strictly valid for all values of $\rho$, but only because they are not as convenient for numerical work when $\rho \ll 1$, have the other expressions been developed.

**E. Potential distribution on electrode surface**

Up to the present the discussion has dealt with potential functions which by construction give no flux across the faces of the disk in which the electrode is imbedded and which have been shown to have very nearly uniform values over large cylindrical envelopes coaxial with the electrodes. In the derivation of these functions it has been explicitly assumed that the flux density, $q$, along the electrode surfaces, was uniform. At the very beginning, however, it was pointed out that physically the electrode must correspond to an equipotential surface, or that over its surface, $\Phi$ was to be constant, as required by the second Eq. (1d). The problem therefore remains of seeing to what extent the potential functions discussed above do satisfy this condition and how they may be made to do so in case they vary considerably over the electrode surface.

In considering the potential distribution over the electrode surface, i.e., for values $\rho = \rho_2$, $w \leqq x$, it is to be remembered that $\rho_2 \ll 1$ for those electrodes that are considered of practical interest from the point of view of such applications as the one to be discussed in another paper. With this in mind one can drop the series in $\rho$ in Eq. (19). The singularities of the coefficients of this series for $w = x$ are only apparent since:

$$\zeta(s, y) - \frac{1}{y^s} = \zeta(s, y + 1)$$
$$= \zeta(s, 1) = 0(1) \text{ for } y = 0; s \geqq 2.$$

Since the whole series development of Eq. (19) is valid only if $\rho < 1 - w - x$ the case $w = x \leqq 1/2$, for which $\zeta(s, 1 - w - x) \leqq (1 - w - x)^{-s}$, will cause no difficulty as these terms are cancelled by the equivalent approach to a singularity of $\log \sin \pi(x + w)$ as $x = w \to 1/2$.

One can therefore safely reduce Eq. (19) to the form:

$$\left.\begin{aligned}\Phi = q\Big[ & 2\log\pi - 2\log\Gamma(w+x)\Gamma(1-w+x) \\ & - \log\sin\pi(w+x)\sin\pi(w-x) + \log\frac{w-x}{w+x} \\ & + \log\frac{w+x+(\rho_2{}^2+(w+x)^2)^{1/2}}{w-x+(\rho_2{}^2+(w-x)^2)^{1/2}}\Big].\end{aligned}\right\} \tag{29}$$

Expanding the logarithmic terms and noting that

$$\log\left\{-b+(\rho_2{}^2+b^2)^{1/2}\right\} = \log\frac{1}{2}\;\frac{\rho_2{}^2}{b}:\frac{\rho_2}{b} \ll 1,$$

one can express $\Phi$ in the following way:

$$\left.\begin{aligned}&x-w \gg \rho_2\text{:} \\ &\quad \Phi = q\Big[2\log\frac{2\pi}{\rho_2} - 2\log\Gamma(x+w)\Gamma(x-w) \\ &\qquad - \log\sin\pi(x+w)\sin\pi(x-w)\Big]; \\ &x-w \leqq \rho_2\text{:} \\ &\quad \Phi = q\Big[\log 2\pi^2 - 2\log\Gamma(x+w)\Gamma(x-w) \\ &\qquad - \log\sin\pi(x+w)\sin\pi(x-w) - \log(x-w) \\ &\qquad - \log\left\{w-x+(\rho_2{}^2+(w-x)^2\right\}\Big]; \\ &x = w\text{:} \\ &\quad \Phi = q\Big[\log\frac{2\pi}{\rho_2} - 2\log\Gamma(2x) - \log\sin 2\pi x\Big] \\ &w > x\text{:} \\ &\quad \Phi = q\Big[2\log\frac{\Gamma(w-x)}{\Gamma(w+x)} + \log\frac{\sin\pi(w-x)}{\sin\pi(w+x)} \\ &\qquad - \log\left\{\tfrac{1}{2}+\tfrac{1}{2}\left(1+\frac{\rho^2}{(w-x)^2}\right)^{1/2}\right\}\Big].\end{aligned}\right\} \tag{29a}$$

It therefore appears that in the limit of very small $\rho_2$ where the terms in Eq. (29a) independent of $\rho_2$ are small compared to $\log\rho_2$, the potential at the bottom of the electrode is but half as large as that along the upper part of the electrode, i.e.,

$$\Phi = 2q\log\frac{2\pi}{\rho_2} \;:\; x-w > \rho_2$$

$$\Phi = q \log \frac{2\pi}{\rho_2} \; : \; x - w = 0.$$

Thus the potential along the electrode is far from constant. The variation of $\Phi$ along the electrode was noted by Samsioe from a numerical calculation for the case where the electrode had but one image—its own reflection in the plane $z=0$—but from the above it appears that in the limit the variation always involves a factor of 2.

The physical reason for this factor of 2 is not difficult to see. Thus for points very close to the electrode, the potential at such points is largely due to the flux elements in their immediate neighborhood both above and below. At the end of the electrode, there are no more flux elements below the end point to contribute to the potential and its magnitude can therefore be only about half as large as that farther up the electrode where there are flux elements on either side of the point to contribute to its potential. This will be shown in section **A,** Part III, to hold even for a variable flux density along the electrode surface.

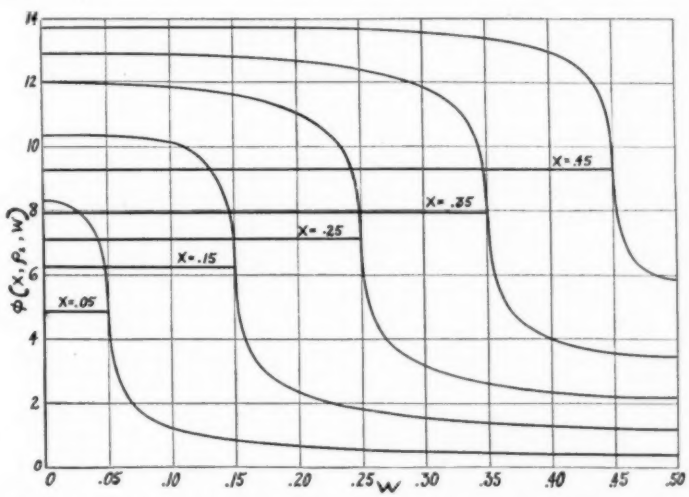



Fig. 5. Potential distribution over surface $\rho_2 = 1/600$, as a function of depth, $w$, from the top of the disk; $x$ = (length of electrode)/2 (disk thickness). Potential function $\Phi$ is that given by Eqs. (29a).

When $\rho_2$ is not so small that the other terms in Eq. (29a) can be neglected, there is still a very considerable variation, although it does not reach the factor of 2. This is shown more clearly in Fig. 5, in which are plotted the functions $\Phi/q$, i.e., the potential for unit flux density, as a function of $w$ for $\rho_2 = 1/600$, for various values of electrode penetration $x$. Thus the curves up to $w = x$,—indicated by the horizontal lines—represent the potential distribution on the electrode surface, the remainder up to $x = 0.5$ being the distribution in the disk just below the electrode, i.e., at the same distance from the axis. These curves show clearly that the assumption of a uniform flux distribution along the electrode surface must be dropped and an attempt to

obtain a uniform potential must be made by varying the flux density along the electrode.

## III. Partially Penetrating Electrodes with Uniform Potentials over Electrode Surface

### A. Potential distribution on electrode surface for variable flux density

It has just been seen that for a constant flux density, $q$, along the electrode surface the potential over the electrode is far from uniform. To eliminate this variation one should in principle consider $q$ in Eq. (17a) or (20) as a function of $\mu$ and determine it in such a way that after integration $\Phi$ will be uniform for the preassigned value of $\rho$. This would lead to an integral equation of the form

$$\left.\begin{aligned} \frac{\Phi}{4} &= \sum_1^\infty n \frac{K_0(2n\pi\rho)}{n\pi} \cos 2n\pi w \int_0^x a(\mu) \\ &\sin 2n\pi\mu d\mu + c \log \frac{2}{\rho}, \quad c = 4\int_0^x \mu a(\mu) d\mu \end{aligned}\right\} \tag{30}*$$

in which $a(\mu)$ is to be determined so that $\Phi$ is constant for $\rho=\rho_2$, $w \leqq x$. The exact solution of this equation would be quite difficult except for very small values of $\rho_2$, in which case the following treatment will indicate the nature of the result.

For this purpose, it will be more convenient to use Eq. (20) directly. It will be assumed that $q$ is everywhere finite and possesses a continuous derivative. Then:

$$\left.\begin{aligned} \Phi = 4\int_0^x q(\mu)d\mu \Big[ &\sum_1^\infty K_0(2\pi n\rho_2) \cos 2\pi n(w-\mu) \\ &+ \cos 2\pi n(w+\mu)\Big\} + \log \frac{2}{\rho_2}\Big]. \end{aligned}\right\} \tag{31}$$

Let:

$$\bar{Q} = \int_0^x q(\mu)d\mu.$$

Then, as the above series is uniformly convergent as long as $\rho_2>0$, $\Phi$ may be written as:

* It is of course appreciated that any function of the form:

$$\Phi = 4A \sum_1^\infty \frac{\cos 2n\pi w K_0(2n\pi\rho)}{K_0(2n\pi\rho_2)} \int_x^{1/2} q(\rho_2, \mu) \cos 2n\pi\mu d\mu + B \log \rho + C$$

will give a strictly constant value for $\Phi$ at $\rho=\rho_2$ $w \leqq x$. However, this value of $\Phi$ does not correspond to an electrode of depth $x$ since the flux $\delta\Phi/\delta\rho$ for $\rho=0$ does not vanish for $w \geqq x$. One has therefore to choose $q(\rho_2, \mu)$ such that $\delta\Phi/\delta\rho=0$ for $w>x$, and then the problem is essentially equivalent to the above in which the *form* of $\Phi$ ensures that $\delta\Phi/\delta\rho=0$ for $w>x$ and the function $a(\mu)$ is to be chosen so that $\Phi$ is constant for $w<x$.

$$\left.\begin{aligned}\Phi = 4\bar{Q}\log\frac{2}{\rho_2} + \frac{2}{\pi}\sum\frac{1}{n}K_0(2\pi n\rho_2)\Big[q(x)\{\sin 2\pi n(x-w)\\ + \sin 2\pi n(x+w)\} - \int_0^x q'(\mu)d\mu\{\sin 2\pi n(\mu-w)\\ + \sin 2\pi n(\mu+w)\}\Big].\end{aligned}\right\} \tag{31a}$$

Now $K_0(\rho)$ has a logarithmic singularity at $\rho=0$. Hence considering $\rho_2$ so small that $\log\rho_2$ is the predominant term in the expansion for $K_0$ it may be replaced by:

$$K_0(2\pi n\rho_2) = \log\frac{2}{\rho_2} \tag{31b}$$

even when $n$ is large, since such terms anyway contribute but little because of the coefficient of $1/n$ in Eq. (31a). $\Phi$ may therefore be rewritten as:

$$\left.\begin{aligned}\Phi = 4\bar{Q}\log\frac{2}{\rho_2} + \frac{2}{\pi}\log\frac{2}{\rho_2}\sum\frac{1}{n}\Big[q(x)\{\sin 2n\pi(x-w)\\ + \sin 2n\pi(x+w)\} - \int_0^x q'(\mu)d\mu\{\sin 2n\pi(\mu-w)\\ + \sin 2n\pi(\mu+w)\}\Big].\end{aligned}\right\} \tag{31c}$$

Noting now that

$$\sum_1^\infty \frac{\sin n\pi y}{n} = \frac{\pi}{2}(1-y):0<y<2 \tag{31d}$$

$\Phi$ may be reduced to:

$$\left.\begin{aligned}\Phi = 4\bar{Q}\log\frac{2}{\rho_2} + 2(1-2x)q(x)\log\frac{2}{\rho_2}\\ - \frac{2}{\pi}\log\frac{2}{\rho_2}\sum\frac{1}{n}\int_0^x q'(\mu)d\mu\{\sin 2n\pi(\mu-w)\\ + \sin 2n\pi(\mu+w)\}.\end{aligned}\right\} \tag{31e}$$

Observing that:

$$\begin{aligned}&\int_0^x q'(\mu)d\mu\sum\frac{1}{n}\sin 2\pi n(\mu-w)\\ &= -\int_0^w q'(\mu)d\mu\sum\frac{1}{n}\sin 2n\pi(w-\mu)\\ &\quad + \int_w^x q'(\mu)d\mu\sum\frac{1}{n}\sin 2n\pi(\mu-w).\end{aligned}$$

and applying again Eq. (31d) one obtains finally,*

$$\Phi = 2q(w) \log \frac{2}{\rho_2} \cdot \tag{32}$$

For $x = w$, the $\sin 2n\pi(x-w)$ terms in Eq. (31c) vanish. The result is then:

$$\left.\begin{aligned} Q = 2 \log \frac{2}{\rho_2} \Bigg[ 2\bar{Q} + q(x)(\tfrac{1}{2} - 2x) - \frac{1}{\pi} \int_0^x q'(\mu) \sum \frac{1}{n} \{ - \sin 2\pi n(x-\mu) \\ + \sin 2n\pi(x+\mu) \} d\mu \Bigg] = q(x) \log \frac{2}{\rho_2} \cdot \end{aligned}\right\} \tag{32a}$$

It thus appears that for such small values of $\rho_2$ that the approximation of Eq. (31b) is justified the potential at a point on the electrode surface is simply proportional to the flux density at that point, the electrode extremity, however, requiring twice the flux density to give the same potential as that at a point on the electrode proper. Hence one may obtain even an arbitrary, distribution of potential on the electrode surface by simply assuming that the flux density $q$ in Eq. (20) is proportional to the required potential distribution, and doubling its value at the bottom of the electrode.

However, as already pointed out, the solution just obtained of making the flux density distribution proportional to the required electrode potential distribution can only be valid if Eq. (31b) is a good approximation, i.e., for extremely small values of $\rho_2$. In practical applications to artesian wells the well radius, $r_2$ is usually about $1/2'$ and the sand thickness $a/2$ is of the order of $50'$, so that $\rho_2 = 1/200$. Now $\log 2/\rho_2 = 5.99$ whereas $K_0(2\pi n\rho_2) < 4.0$ for even $n = 1$. Hence it is clear that Eq. (31b) is not a valid approximation for cases of practical interest from the point of view of water flow into artesian wells. Before, however, presenting a method that will be applicable, it is of interest to see what a uniformly increasing flux density would give for the electrode potential.

### B. Potential distribution for a linearly increasing flux density

Since the case of a uniform flux has already been discussed it will be perfectly general to take

$$q(\mu) = \mu. \tag{33}$$

Although the simplest formula would be obtained by applying Eq. (33) to (20), it will be better for numerical purposes to apply it to (17a) since it is particularly adapted for computation when $\rho_2$ is very small.

Denoting, then, the coefficient of $q$ in the right side of Eq. (17a) by $d\Phi_0$ it is clear that for the present case:

$$\Phi = \int_0^x \mu d\Phi_0 = x\Phi_0(\rho, w, x) - \int_0^x \Phi_0(\rho, w, \mu) d\mu, \tag{34}$$

* If the same process is carried through for $w > x$ it will be found that $\Phi = 0$. This simply means that $q(\mu) = 0$, $\mu > x$ as the form of (31) indicates. To get the true value of $\Phi$ for $w > x$, $\rho_2 < < 1$ one must therefore include the constant terms in the expansion for $K_0$.

where by Eq. (19)

$$\left.\begin{aligned}\Phi_0(\rho, w, \mu) &= \log\frac{w+\mu+(\rho^2+(w+\mu)^2)^{1/2}}{w-\mu+(\rho^2+(w-\mu)^2)^{1/2}}\\ &- 2\log\Gamma(w+\mu)\Gamma(1-w+\mu) - \log\sin\pi(w+\mu)\sin\pi(w-\mu)\\ &+ \log\frac{w-\mu}{w+\mu} - \frac{1}{4}\rho^2\left\{\zeta(2,1-w-\mu) - \zeta(2,1-w+\mu)\right.\\ &\left.+ \zeta(2,w-\mu) - \zeta(2,w+\mu) + \frac{1}{(w+\mu)^2} - \frac{1}{(w-\mu)^2}\right\}\\ &+ \frac{3}{32}\rho^4\left\{\zeta(4,1-w-\mu) - \zeta(4,1-w+\mu) + \zeta(4,w-\mu)\right.\\ &\left.- \zeta(4,w+\mu) + \frac{1}{(w+\mu)^4} - \frac{1}{(w-\mu)^4}\right\}.\end{aligned}\right\} \quad (34a)$$

Evaluating the integral in Eq. (34) it is found that:

$$\left.\begin{aligned}\int_0^x \Phi_0(\rho, w, \mu)d\mu &= (w+x)\sinh^{-1}\frac{w+x}{\rho}\\ &+ (w-x)\sinh^{-1}\frac{w-x}{\rho} - 2w\sinh^{-1}\frac{w}{\rho} - (\rho^2+(w+x)^2)^{1/2}\\ &- (\rho^2+(w-x)^2)^{1/2} + 2(\rho^2+w^2)^{1/2} - (w+x)\log(w+x)\\ &- (w-x)\log(w-x) + 2w\log w + \tfrac{1}{4}\rho^2\left\{\psi(1-w-x)\right.\\ &\left.- 2\psi(1-w) + \psi(1-w+x) + \psi(w-x) - 2\psi(w)\right.\\ &\left.+ \psi(w+x) + \frac{1}{w+x} + \frac{1}{w-x} - \frac{2}{w}\right\} + \frac{1}{32}\rho^4\left\{\zeta(3,1-w-x)\right.\\ &\left.- 2\zeta(3,1-w) + \zeta(3,1-w+x) + \zeta(3,w-x)\right.\\ &\left.- 2\zeta(3,w) + \zeta(3,w+x) - \frac{1}{(w+x)^3} - \frac{1}{(w-x)^3} + \frac{2}{w^3}\right\}\\ &+ I(\rho, w, x),\end{aligned}\right\} \quad (34b)$$

where:

$$\left.\begin{aligned}-I(\rho, w, x) &= \int_0^x \left\{2\log\Gamma(w+\mu)\Gamma(1-w+\mu)\right.\\ &\left.+ \log\sin\pi(w+\mu)\sin\pi(w-\mu)\right\}d\mu.\end{aligned}\right\} \quad (34c)$$

Although it is difficult to integrate the individual terms in Eq. (34c) they may be paired and integrated in the following way. Thus, from the theory of gamma-functions it is known that:

$$2 \log \Gamma(y) + \log \sin \pi y = \log \pi + 2 \sum_{1}^{\infty} n a_n \sin 2n\pi y; \; 0 < y < 1, \quad (35)^*$$

where:

$$a_n = \frac{1}{n\pi}(\delta + \log 2\pi n), \quad (35a)$$

$$\delta = 0.5772.$$

Hence

$$\left.\begin{aligned} -I &= x \log \pi - \sum_{1}^{\infty} \frac{a_n}{n\pi}\{\cos 2n\pi(w + x) - \cos 2n\pi w\} \\ &+ \int_0^x \{2 \log \Gamma(1 - w + \mu) + \log \sin \pi(w - \mu)\} d\mu. \end{aligned}\right\} \quad (35b)$$

Since $1-w+\mu>1$ for $\mu>w$, and (35) is valid only for $y<1$,—I must be evaluated separately for the cases of $w>x$ and $w<x$. For the latter case the integral in Eq. (35b) must be split up as:

$$\int_0^x = \int_0^w + \int_w^x.$$

With this precaution one finds:

$w > x$:

$$-I = 2x \log \pi - 2 \sum_{1}^{\infty} \frac{a_n}{n\pi} \cos 2n\pi w\{\cos 2n\pi x - 1\}; \quad (35c)$$

$w \leqq x$:

$$\left.\begin{aligned} -I &= 2x \log \pi - 2 \sum_{1}^{\infty} \frac{a_n}{n\pi} \cos 2n\pi w\{\cos 2n\pi x - 1\} \\ &+ 2(x - w)\{\log (x - w) - 1\} + (x - w) \log (-1). \end{aligned}\right\} \quad (35d)$$

It is to be noted that the term $(x-w) \log (-1)$ in (35d) is just right to cancel the equivalent imaginary term in $(x-w) \log (w-x)$ in Eq. (34b) for $w<x$. Hence $\Phi$ remains real in all cases, as of course it should.

By means of the above equations the potential distribution over the electrode surface, i.e., for $\rho_2=0.001$ in the present case, was computed for a 50 percent electrode penetration, i.e., for $x=0.25$. The result is indicated in curve I, Fig. 6.

It will be noticed that the potential in curve I follows the flux density for about 80 percent of the length of the electrode and after that the end effect—in which the potential at the electrode extremity falls to about half of its normal value—begins to be significant. Strangely enough, however, when the proper uniform flux density is added to the linearly increasing flux one

* It may be of interest to note that by using Eq. (35), after expanding $K_0(2n\pi\rho)$ to terms in $\rho^2$, one may show in detail the exact equivalence of Eq. (19) and (21) up to terms in $\rho^2$.

obtains an almost exactly linear variation in the result potential. This is indicated in curve II of Fig. 6.

In a manner similar to that used above other functional forms could be tried for $q(\mu)$ and the corresponding potential distributions could be computed and plotted. It is clear, however, that one is not likely to find a suitable

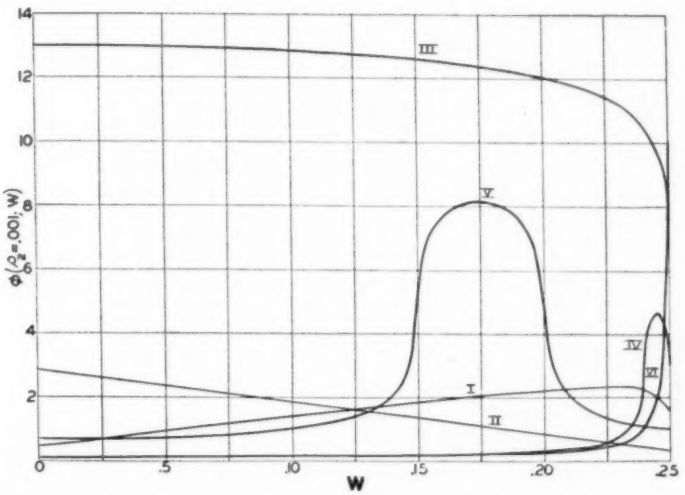



Fig. 6. Potential distributions over the surface of a 50 percent penetrating electrode of radius $\rho_2=0.001$, for various types of flux distribution. Curve I: potential function of Eq. (34) due to a linearly increasing flux density. Curve II: potential function of Eqs. (34) and (29) due to a flux density $q(\mu)=-1/4+\mu$. Curve III: potential function $\Phi_0(0.25, w)$ of Eq. (36) due to a uniform unit flux density. Curve IV: potential function $\Delta\Phi(0.24, 0.25, w)$ of Eq. (36a) due to a unit flux density between $w=0.24$ and $w=0.25$; Curve V: potential function $\Delta\Phi(0.15, 0.20, w)$; Curve VI: potential function $\Phi_p(0.25, w)$ of Eq. (36b) due to a 0.01 unit source at $w=0.25$.

form for $q(\mu)$ unless complicated or "trick" functions are resorted to. It therefore has appeared best to use the more direct, though less elegant, method of superposing flux elements at various places along the electrode axis and adjusting their strengths so that the resultant potential over the electrode surface is approximately uniform. This method will now be described in more detail.

### C. Potential distribution due to a superposition of flux elements

Since the discussion of this section will be confined essentially to the potential for very small values of $\rho_2$—i.e., the potentials on the electrode surface—the terms in $\rho_2^2$ will be definitely dropped from all the expressions and the potential functions used will be those given in Eqs. (29a).

For definiteness it is convenient to introduce the following notation:

$$\left.\begin{aligned}\Phi_0(x, w) &= 2 \log \frac{2\pi}{\rho_2} - 2 \log \Gamma(x + w)\Gamma(x - w) \\ &\quad - \log \sin \pi(x + w) \sin \pi(x - w)\end{aligned}\right\}. \tag{36}$$

$$\Delta\Phi(x_1, x_2, w) = \Phi_0(x_1, w) - \Phi_0(x_2, w). \tag{36a}$$

$$\left.\begin{aligned}\Phi_p(x, w) = &- 2\psi(1 - w + x) - 2\psi(1 + w + x)\\ &- \pi \operatorname{ctn} \pi(x + w) - \pi \operatorname{ctn} \pi(x - w)\\ &+ \frac{1}{x + w} + \frac{1}{x - w} + \frac{1}{(\rho_2^2 + (w + x)^2)^{1/2}}\\ &+ \frac{1}{(\rho_2^2 + (x - w)^2)^{1/2}}.\end{aligned}\right\} \tag{36b}$$

Physically, the above equations have the following meaning. $\Phi_0(x, w)$ is the potential at $(\rho_2, w)$ due to a uniform unit flux distribution on the $z$-axis and extending from $\mu = 0$ to $\mu = x$. It is to be understood that the expression given for $\Phi$ is to be changed as in Eq. (29a) if $x - w \sim \rho_2$ or $x = w$. $\Delta\Phi(x_1, x_2, w)$ then represents the potential due to a unit flux element between $\mu = x_1$ and $\mu = x_2$. Finally, $\Phi_p(x, w)$ is the potential due to a unit source at $\mu = x$. It is obtained most conveniently from Eq. (17a) by setting $qd\mu = 1$, $\mu = x$. Of course, all the functions $\Phi_0$, $\Delta\Phi$, and $\Phi_p$ satisfy Laplace's equation and give zero flux over $w = 0, 1/2$.

The method of superposition is simply that of superposing various combinations of the functions in Eq. (36) such that the resultant gives an approximately uniform value for $w \leqq x$. It is a rather tedious task and takes considerable time but after a few trials one's experience will suggest proper "guesses" for the proper strengths and distributions of the functions.

To indicate more clearly the nature of these adjustments and the meaning of the functions, plots of some representative cases, for $\rho_2 = 0.001$, are given in Fig. 6. Although finally, the resultant potential function is to be expressed as a sum of terms as $\Phi_0(x, w)$ for different values of $x$, it is more convenient at first to use the functions $\Delta\Phi$, since from their form one can more easily tell just how to build up the potential where it may deviate from its average value. The forms of the curves for $\Delta\Phi$ as given in curves IV and V also show clearly that due to the "end effects" one cannot build up the potential near the end of the electrode, by finite flux elements alone. Rather one has to resort to the use of the point source placed at the electrode extremity. Of course, even then one does not obtain a strictly uniform potential over the electrode surface; however, a potential distribution whose maximum differences do not exceed 2 percent of the average potential can be obtained without very much difficulty and for practical purposes this is of course fully sufficient.*

As a practical aid, tables were constructed giving $\log \Gamma(y)$, $\log \sin \pi y$ and $\pi \operatorname{ctn} \pi y$ for $0 \leqq y \leqq 1$ the values of $y$ increasing by intervals of 0.005. From these, tables of $\Phi_0(x, w)$ and $\Phi_p(x, w)$ were computed for various values of $x$ and these served to give the $\Phi\Delta$.

* Although the above analysis has said nothing of the bottom of the electrode it really corresponds to an almost hemispherical cap. From Eq. (29) it follows that $\Phi_0(x, \rho_2, x) = \Phi_0(x, 0, x + \rho_2) + \log 2$. The bottom of the electrode is therefore not exactly spherical, but is slightly flattened.

In Fig. 7 are plotted the resultant values of $\Phi$ over the electrode surface and on its geometrical extension to the bottom of the disk, i.e., to $w=0.5$ for various values of $x$. Their analytical representations are given by the following equations in which are included the complete penetration and zero penetration cases. The functions $\Phi_0$ and $\Phi_p$ in the equations below are really the

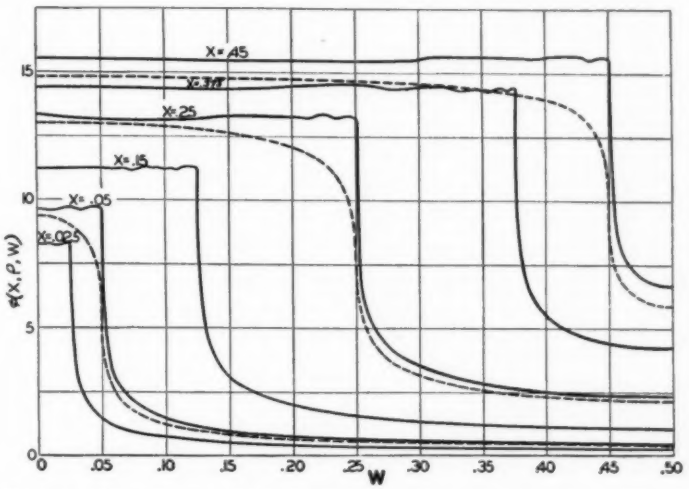



Fig. 7. Curves giving the potential distributions over the electrode surfaces of radius $\rho_2=0.001$, after adjustment of the flux distribution according to equations shown below. Dotted curves represent the unadjusted distributions.

functions of Eq. (19) and (17a) which reduce to the corresponding functions of Eq. (36) and (36b) when $\rho$ is equal to $\rho_2$, the effective electrode radius. The coefficients have been adjusted so that the resultant $\Phi$ reduces to the uniform value $\Phi_c$ when $\rho_2=0.001$. The constant $\Phi_c$ has then been added to the expressions so as to make $\Phi$ vanish on the electrode surface. The results are:

$x=0.50$: electrode penetration $=100\%$.

$$\Phi(0.50, \rho, w) = 15.2018 - 2\log\frac{2}{\rho}$$

$x=0.45$: electrode penetration $=90\%$.

$$\begin{aligned}\Phi(0.45, \rho, w) = 15.62 - \{&1.433\Phi_0(0.45, \rho, w) - 0.233\Phi_0(0.44, \rho, w)\\ &- 0.05\Phi_0(0.42, \rho, w) - 0.05\Phi_0(0.40, \rho, w) - 0.02\Phi_0(0.375, \rho, w)\\ &- 0.04\Phi_0(0.30, \rho, w) + 0.0038\Phi_p(0.45, \rho, w)\}.\end{aligned}$$

$x=0.375$: electrode penetration $=75\%$.

$$\begin{aligned}\Phi(0.375, \rho, w) = 14.41 - \{&1.35\Phi_0(0.375, \rho, w) - 0.18\Phi_0(0.36, \rho, w)\\ &- 0.07\Phi_0(0.34, \rho, w) - 0.05\Phi_0(0.30, \rho, w) - 0.05\Phi_0(0.20, \rho, w)\\ &+ 0.0042\Phi_p(0.375, \rho, w) + 0.0002\Phi_p(0.37, \rho, w)\}.\end{aligned}$$

$x = 0.25$: electrode penetration = 50%.

$$\Phi(0.25, \rho, w) = 13.28 - \{1.45\Phi_0(.25, \rho, w) - 0.24\Phi_0(0.24, \rho, w) - 0.10\Phi_0(0.22, \rho, w) - 0.031\Phi_0(0.20, \rho, w) - 0.07\Phi_0(.15, \rho, w) + 0.0038\Phi_p(0.25, \rho, w)\}.$$

$x = 0.125$: electrode penetration = 25%.

$$\Phi(.125, \rho, w) = 11.25 - \{1.34\Phi_0(0.125, \rho, w) - 0.22\Phi_0(0.115, \rho, w) - 0.07\Phi_0(0.10, \rho, w) - 0.05\Phi_0(0.08, \rho, w) - 0.03\Phi_0(0.05, \rho, w) + 0.00354\Phi_p(0.125, \rho, w)\}.$$

$x = 0.05$: electrode penetration = 10%.

$$\Phi(0.05, \rho, w) = 9.68 - \{1.25\Phi_0(0.05, \rho, w) - 0.19\Phi_0(0.04, \rho, w) - 0.06\Phi_0(0.02, \rho, w) + 0.00333\Phi_p(.05, \rho, w)\}.$$

$x = 0.025$: electrode penetration = 5%.

$$\Phi(0.025, \rho, w) = 8.30 - \{1.30\Phi_0(.025, \rho, w) - 0.26\Phi_0(0.02, \rho, w) - 0.04\Phi_0(0.01, \rho, w) + 0.0028\Phi_p(0.025, \rho, w)\}.$$

$x = 0.001$: electrode penetration = 0.2%—"nonpenetrating electrode".

$$\Phi(0.001, \rho, w) = 2002.308 - \Phi_p(0, \rho, w).$$

To illustrate more clearly the effect of the superposition, the original unadjusted potential functions $\Phi_0(x, \rho, w)$ have been plotted in Fig. 7 for a few

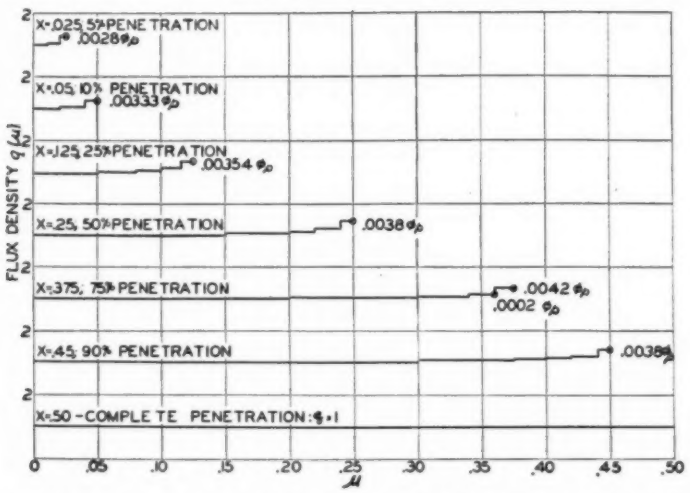



Fig. 8. Flux distributions over electrode surfaces required to give the potential distributions of Fig. 7. $a_p\Phi_p$ represents a point source of strength $a_p$.

values of $x$ for $\rho = \rho_2 = 0.001$. These are the broken curves. From these it appears that near the bottom of the disk the potential due to a uniform flux along the electrode and that due to a uniform potential over the electrode

surface run closely parallel to each other. However, near the electrode extremities the potentials fall off much more steeply for the adjusted potential distribution. This of course corresponds to the excess flux density near the bottom of the electrode which was added to make the potential uniform over the electrode surface.

Since the physical meaning of the superposition method used above is that of the superposition of flux elements, a plot of the flux distributions may be of interest. These distributions have been plotted in Fig. 8 for $\rho_2 = 0.001$, and various values of $x$. As in no case does the flux density exceed 2, the scale has been shifted for the various values of $x$. The discontinuous curves of Fig. 8 are clearly the approximations implied by the superposition method, of the continuous flux distributions which would give a perfectly uniform potential over the electrode surface. By drawing smooth curves through the discontinuous segments, one should get a still closer approximation to the actual continuous flux distribution over the electrode surface.

These curves show that not only does the flux always require a concentrated point sink at the electrode extremity, which must be added to the more continuous distribution along the electrode proper, but also this latter distribution appears to be too concentrated near the electrode extremity to be represented by any very simple continuous function. One may note, too, that, as is to be expected, the relative steepness of the rise of flux near the electrode extremity decreases with increasing penetration. This, of course, corresponds to the decreasing relative importance of the nonradial part of the flux with increasing penetration.

Before closing this discussion it may be of interest to show more clearly that the above superposition method is essentially nothing more than the direct approximate solution of the integral Eq. (30). Thus, it will be clear that the superposition of the functions $\Phi_0$, $\Delta\Phi$, and $\Phi_p$ simply amounts to writing the resultant potential functions, except for a constant, in the form:

$$\Phi = \sum a_m \Phi_0(x_m, \rho, w)$$

which according to earlier formulae may be expanded as:

$$\frac{\Phi}{4} = \sum_1^\infty n \frac{K_0(2n\pi\rho)}{n\pi} \cos 2n\pi w \left\{ \sum a_m \sin 2n\pi x_m + 2n\pi a_p \cos 2n\pi x_p \right\} + (a_p + \sum x_m a_m) \log \frac{2}{\rho},$$

where $a_p$ represents the strength of the point source at $x = x_p$.

Returning now to the integral Eq. (30) it will be noted that the most direct analytical approximation method that could be used to solve it would be the following. First one would replace the integral $\int_0^x a(y) \sin 2n\pi y dy$ by $\sum_m a_m \sin 2n\pi x_m$, $x_m \leqq x$, with evenly spaced $x_m$ and then after choosing $\rho_2$ one would set the right side of Eq. (30) equal to a constant for each of a set of values of $w$, not exceeding in number the number of values of $x_m$. These would give a linear system of equations for the coefficients $a_m$ which could be solved by the usual algebraic processes.

Now as seen by comparing Eqs. (30) and (31), $a(\mu) \sim q'(\mu)$. Hence as $q(\mu)$ varies but slightly for small values of $w$, the $a_m$ for small values of $x_m$ would be small and can be neglected to begin with. Further, it appears that due to the very rapid drop in $\Phi_0$ for $w = x$ anything but a very close spacing of the $x_m$ would give considerable oscillations in $\Phi$ between the $x_m$. From these considerations it is seen that the natural way of treating the above proposed linear equations in $a_m$ would be to deliberately drop the terms $a_m \sin 2n\pi x_m$ for small $x_m$ and then avoid the close spacing near $x_m = x$ by adding a point source at $x_m = x$. But this procedure is clearly essentially identical with the method actually used and expressed by the above formula for $\Phi$. It is seen therefore that the present superposition method is really a direct method of solving the integral Eq. (30) in which, however, the resulting algebraic equations were simplified and solved graphically by applying the available a priori information as to the nature of the functions $\Phi_0(x_m, \rho, w)$.

**D. Potential distribution for large values of $\rho$ with uniform electrode potential**

To see how $\Phi$ behaves when $\rho$ is large, it should be recalled from the last section that for practical purposes $\Phi$ can be represented by the expansion:

$$\Phi(x, \rho, w) = \Phi_c - \sum a_m \Phi_0(x_m, \rho, w) : x_m \leqq x, \tag{37}$$

where $\Phi_c$ is the constant potential over the electrode surface, the $\Phi_0\ (x_m, \rho, w)$ are the potential functions of Eq. (19) for $x = x_m$—representing the potential due to a uniform unit flux density extending from $\mu = 0$ to $\mu = x_m$—and $a_m$ is an effective "strength" associated with $\Phi_0(x_m, \rho, w)$ i.e., it gives the numerical flux density for the particular flux element extending from $\mu = 0$ to $\mu = x_m$. It is to be understood, further, that at least one of the terms $\Phi_0(x_m, \rho, w)$ is to be the potential function for a point source, usually situated at $x_m = x$, the potential of which is given by Eq. (17a) with $qd\mu = 1$, $\mu = x$.

Now as already pointed out the analytical equivalents of the functions $\Phi_0(x_m, \rho, w)$, which are particularly suitable for use at large values of $\rho$, are given by Eq. (21): this means that:

$$\Phi_0(x_m, \rho, w) = 4\left[\frac{1}{\pi} \sum_1^\infty n \frac{1}{n} K_0(2n\pi\rho) \cos 2n\pi w \sin 2n\pi x_m + x_m \log \frac{2}{\rho}\right]. \tag{37a}$$

Since, as has been seen in the earlier sections, the terms in the summation are very small for values of $\rho \geqq 1$, one has immediately as the approximation to $\Phi_0(x_m, \rho, w)$, valid for large $\rho$:

$$\Phi_0(x_m, \rho, w) = 4x_m \log \frac{2}{\rho}. \tag{38}$$

Similarly:

$$\Phi_\rho(x_m, \rho, w) = 4 \log \frac{2}{\rho}. \tag{38a}$$

One can therefore rewrite Eq. (37) in the form:

$$\Phi = \Phi_c - 4 \sum a_m x_m \log \frac{2}{\rho} : \rho \geqq 1 \tag{39}$$

where $x_m = 1$ for the $a_m$ referring to $\Phi_p$.

It is thus seen that for large values of $\rho$, the potential is for all practical purposes independent of the depth and varies logarithmically with the radial distance from the electrode. This is more clearly seen from Fig. 9 in which are plotted a number of equipotential curves for $x = 0.25$, $\rho_2 = 0.001$, $\rho_1 = 2$. The potential functions used are really the $\Sigma a_m \Phi_0$ of Eq. (36) since the term $\Phi_c$ adds nothing to the nature of the distribution.

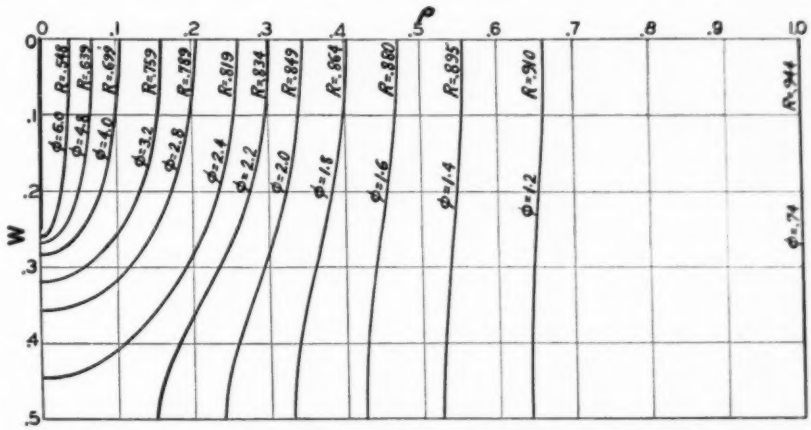



Fig. 9. Equipotential curves in a disk with a 50 percent penetrating electrode. $\Phi$ is the function given by Eq. (37). $R$ = fraction of total potential drop across the disk.

Again from Eq. (39) one may get a convenient check with regard to the total flux passing through the system: As in Eq. (22), this is given by:

$$Q = -2\pi a \int_0^{1/2} \rho \frac{\partial \Phi}{\partial \rho} dw = 4\pi a \sum a_m x_m. \tag{40}$$

This result is quite rigorous, since even if the exact Eqs. (37a) had been used the trigonometric terms would have vanished on integration.

### E. The resistance of the disks as a function of the penetration

Recalling that Eq. (39) gives the potential drop between the electrode and external boundary, at $\rho_1$, the resistance of the system can be obtained as in Part I by dividing $\Delta\Phi$, of $\Phi$ of Eq. (39), by the total flux $Q$, as given in Eq. (40). Thus:

$$\Omega = \frac{\Delta\Phi}{Q} = \frac{1}{\pi a}\left[\frac{\Phi_c}{4\sum x_m a_m} - \log \frac{2}{\rho_1}\right]. \tag{41}$$

For purposes of plotting, however, it is more convenient to use the dimensionless function

$$\frac{a\Omega}{2} = \frac{1}{2\pi}\left[\frac{\Phi_c}{4\sum x_m a_m} - \log\frac{2}{\rho_1}\right] \tag{41a}$$

which gives the resistance of the system for unit thickness of the disk. In Eq. (41a) the $\Phi_c$ and $4\Sigma x_m a_m$ are to be obtained from the expressions in Section C, even for the cases $x = 0.50$ and $x = \rho_2$. Assuming $\rho_1/\rho_2 = r_1/r_2 = 2000$, the curves in Fig. 10 were plotted from Eq. (41a). Logarithmic paper was used because of the large variation in $\Omega a/2$. The difference between the curves I and II is clearly a measure of the dependence of the resistance of the disk upon its absolute thickness. Thus it appears that the effect of the absolute thickness decreases with increasing penetration.

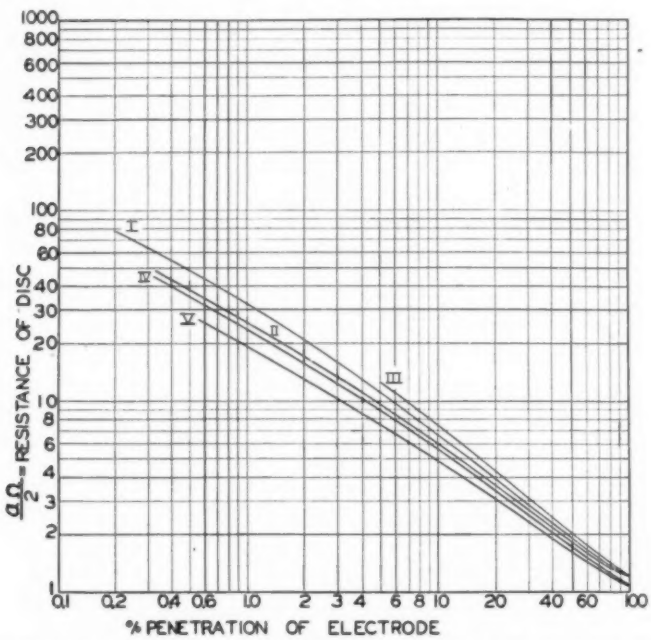



Fig. 10. Resistance of the disk as a function of the electrode penetration. $a\,\Omega/2$ = resistance of disk per unit disk thickness. Curve I: $\rho_2 = 0.001$, $\rho_1/\rho_2 = 2000$ with electrode potential uniform, as given by Eq. (41a). Curve II: $\rho_2 = 1/600$; $\rho_1/\rho_2 = 2000$ with electrode potential uniform, Curve III: $\rho_2 = 0.001$, $\rho_1/\rho_2 = 2000$ with uniform flux density over electrode (Eq. (42)); crosses represent values computed by Eq. (43). Curve IV: $\rho_2 = 1/500$, $\rho_1/\rho_2 = 1000$, (computed by Eq. (43)). Curve V: $\rho_2 = 1/300$, $\rho_1/\rho_2 = 1000$, (computed by Eq. (43)).

It will be seen that after a penetration of about 5 percent curves in Fig. 10 are approximately linear with a slope a little less than one. However, the marked curvature appearing for penetrations close to 100 percent is quite real and its existence has, in fact, been verified experimentally.* It is of interest, too, that the resistances for the "nonpenetrating electrodes" given by the 2 points at the extremities of curves I and II fall, without adjust-

* The writer is indebted to Mr. R. D. Wyckoff and Mr. W. O. Smith for these results which were obtained by high frequency experiments performed with a large cylindrical tank containing a $CuSO_4$ solution.

ment, on the natural extrapolations of the main parts of the curves. There are thus no discontinuities or sharp curvatures, so that interpolation for any intermediate points can be made with but little uncertainty.

Since these computations of the resistance by Eq. (41a) requires the preliminary tedious determination of $\Phi_c$ and the constants $x_m$ for each electrode radius or disk thickness, it is of some interest to develop an approximate formula that will avoid as much of this labor as possible. Since it is readily shown that to abandon the infinite set of images is a far too crude approximation, the only other reasonable approximation is to see what can be done with a uniform flux density along the electrode, as developed in Part II. In this case Eq. (41a) would reduce to:

$$\frac{a\Omega}{2} = \frac{1}{2\pi}\left[\frac{\Phi_c}{4x} - \log\frac{2}{\rho_1}\right]. \tag{41b}$$

However, $\Phi_c$ which represents the electrode potential must still be determined. Since the potential over the electrode surface is not constant when the flux density is uniform $\Phi_c$ should correspond to an average of the surface potential. As a first approximation, however, one may assume $\Phi_c$ to be the potential at the top of the electrode, which by Eq. (36) is:

$$\Phi_c = \Phi_0(x, 0) = 2\log\frac{2\pi}{\rho_2} - 4\log\Gamma(x) - 2\log\sin\pi x.$$

Eq. (41b) may then be written as:

$$\frac{a\Omega}{2} = \frac{1}{2\pi}\left[\frac{1}{2x}\log\frac{2x}{\rho_2} + \frac{1}{2x}\log\frac{\Gamma(1-x)}{\Gamma(1+x)} - \log\frac{2}{\rho_1}\right]. \tag{42}$$

This is plotted in Fig. 10 for $\rho_2 = 0.001$ and $\rho_1 = 2.0$ as a function of $x$, as curve III. It will be seen that the $a\Omega/2$ of curve III are almost uniformly 10 percent higher than the correct values given by curve I. Although this agreement would be good enough for some purposes it will be presently seen that a slight change in the assumption for $\Phi_c$ will improve the approximation considerably. But before leaving Eq. (42) it is of interest to compare it with the corresponding result of Samsioe who also assumed a uniform flux density and took $\Phi_c$ as the potential at the top of the electrode. In the present notation Samisoe's result is equivalent to:

$$\frac{a\Omega}{2} = \frac{1}{2\pi}\left[\frac{1}{2x}\log\frac{2x}{\rho_2} + \frac{1}{2x}\log\left(\frac{1+1.44x}{1-1.44x}\right)^{0.4} - \log\frac{2}{\rho_1}\right]. \tag{42a}$$

Since a numerical check easily shows that:

$$\left(\frac{1+1.44x}{1-1.44x}\right)^{0.4} = \frac{\Gamma(1-x)}{\Gamma(1+x)}$$

to a high degree of accuracy, it is clear that the present treatment gives the same result as Samsioe's for the net resistance of the system under the condition of uniform flux density along the electrode surface.

Returning now to the above assumption that $\Phi_e = \Phi_0(x, o)$ it is evident that it would be more reasonable to choose for $\Phi_e$ the value of $\Phi$ at some definite position down along the electrode. By trial it has been found that if $\Phi_e$ is taken as the value of $\Phi_0$ at $w = 3/4x$ the resistances computed according to Eq. (41b) agrees with the values computed for Eq. (41a) to within less than 1/2 percent. The agreement may be seen better from Fig. 10, where the crosses are the values computed as stated above, i.e., from the formula:

$$\frac{a\Omega}{2} = \frac{1}{2\pi}\left[\frac{\Phi_0(x,\ w = \frac{3}{4}x,\ \rho_2)}{4x} - \log\frac{2}{\rho_1}\right], \tag{43}$$

or, in detail

$$\left.\begin{aligned}\frac{a\Omega}{2} = \frac{1}{2\pi}\Bigg[\frac{1}{4x}\Bigg\{2\log\frac{2\pi}{\rho_2} - 2\log\Gamma(1.75x)\Gamma(0.25x)\\ - \log\sin 1.75\pi x \sin 0.25\pi x\Bigg\} - \log\frac{2}{\rho_1}\Bigg].\end{aligned}\right\} \tag{43a}$$

By the use of this formula one is able to avoid the tedious work of adjusting flux elements so as to obtain a strictly uniform electrode potential, when computing the resistance of the system. Thus with this expression curves IV and V of Fig. 10 were computed. These show the variation of $a\Omega/2$ with the electrode radius since they correspond to curves I and II except that the electrode radius has been doubled. As the curves show, the effect is not large, and as is to be expected $a\Omega/2$ varies with the electrode radius logarithmically, as for pure radial flow, for large penetrations, and approaches an inverse first power variation for the small penetrations.

It is to be understood that if one is interested in the details of the potential distribution the above approximation has no meaning and one must use the more lengthy functions of section **C**. It should be pointed out, also, that since the resistance of the system is inversely proportional to the thickness of the disk, the resistance for a disk of thickness $a/2$ will be given by the ordinates in Fig. 10, *divided* by $a/2$. As to the units it is clear that since $a\Omega$ is dimensionless, the units of $\Omega$ will correspond to those of the length $a$ and the specific resistance, which has been taken as unity in all the above considerations.

The above discussion would sufficiently complete the analysis from the electrical point of view. However, from the point of view of fluid flow through a sand, one is usually interested in the flux or production passing through the system for unit potential or pressure difference. That is, the main interest is in the reciprocal of the resistance. It will therefore be useful to reinterpret some of the above results from this point of view.

Thus the flux per unit potential difference or production per unit pressure difference and per unit sand thickness is given by Eq. (41) as:

$$\frac{2Q}{\Delta\Phi a} = \frac{2}{a\Omega} = \frac{2\pi}{\dfrac{4\sum x_m a_m}{\Phi_e} - \log\dfrac{2}{\rho_1}}. \tag{44}$$

In Fig. 11 the quantity $2Q/a\Delta\Phi$ is plotted for several values of $\rho_2$ with $\rho_1/\rho_2 = 2000$. For $\rho_2 - 1/600, 1/1000$ Eq. (44) was used. For $\rho_2 = 1/400, 1/1600$, the reciprocal of Eq. (43) was used, i.e.,:

$$\frac{2Q}{a\Delta\Phi} = \frac{2\pi}{\dfrac{\Phi_0(x,\ w = 0.75x,\ \rho_2)}{4x} - \log\dfrac{2}{\rho_1}}. \tag{44a}$$

Again the differences between the various curves indicate that the flux for various penetrations is not strictly proportional to the absolute disk thickness. Since for a given electrode or well radius the sand thickness is inversely

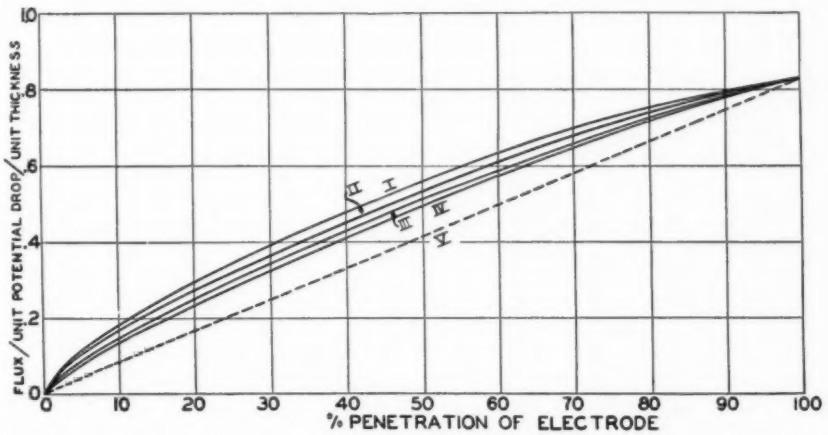



Fig. 11. Curves showing the total flux through the disk for unit potential difference and unit disk thickness, for various disk thicknesses as a function of the electrode penetration. $\rho_1/\rho_2 = 2000$. Permeability or conductivity $= 1.0$. Curve I: $\rho_2 = 1/400$; Curve II: $\rho_2 = 1/600$; Curve III: $\rho_2 = 1/1000$; Curve IV: $\rho_2 = 1/1600$. Curve V: Flux if the flow were purely radial.

proportional to $\rho_2$, the fact that the curves for higher $\rho_2$ lie higher in Fig. 11 shows that the flux per unit disk thickness is greater for the smaller thickness, or that the flux for a given percent of penetration does not increase quite as fast as the disk thickness.

The dotted line in Fig. 11 is the flux that would pass through the system if the flow into the electrode were completely radial for all penetrations. The difference of about 20 percent between this line and the curves above it, shows at once that the actual flow system cannot be considered as a superposition of a radial flow into the electrode proper and a semispherical flow into the electrode extremity, since the contribution due to the latter is only about 2.5 percent of the radial flux. Rather, the flux distribution even appreciably above the electrode extremity is distorted from a radial character and many of the flow lines from a depth below that of the electrode come up and end on the electrode proper.

In order that the results of the above analysis be as general as possible, no attempt has been made to put in absolute numerical values for any of

the dimensional quantities. When such specific results are desired they can be readily obtained from those presented by the multiplication or divison by the proper coefficient. Thus with regard to Fig. 11, if the actual flux is desired for a given pressure differential and disk thickness one need only multiply the numerical ordinates of Fig. 11 by the given pressure differential and thickness, these being expressed, however, in the same units as the permeability, which again has been taken as unity.

The writer wishes to thank Mr. R. D. Wyckoff and Dr. P. D. Foote of the Gulf Research Laboratory for the opportunity of discussing with them questions arising in the preparation of this paper.

# Chemical Treatment of Rotary Drilling Fluids*

By H. C. LAWTON, H. A. AMBROSE, and A. G. LOOMIS
*Mellon Institute of Industrial Research and Gulf Research Laboratory, Pittsburgh, Pa.*

(Received March 11, 1932)

The desirable properties of rotary drilling fluid in the technique of boring oil wells are considered. An analysis of the problem of elimination of cuttings by differential settling with reclamation of the colloidal content of clay dispersions shows that viscosity should be lowered as much as possible when the density is held constant. It is demonstrated that lowering of viscosity results in a small lowering of pump pressure, even though the velocity of the mud-fluid is increased to maintain constant carrying capacity for cuttings. Several chemicals have been found that cause remarkable lowering of the viscosity of clay drilling fluids. Suitable adjustment of the concentration of caustic soda in these chemical reagents imparts desirable thixotropic properties to the clay dispersion. The theoretical results of this study are applied to some of the practical problems of oil-well drilling.

## I. INTRODUCTION

THERE are two standardized methods now in general use for boring oil wells. One method of drilling, known as cable-tool drilling and described as a "standard rig" in oil-field parlance, uses a heavy drilling bit attached to a wire cable. By a suitable mechanism, causing withdrawal and release of the cable, the bit is alternately raised and dropped, striking hammer blows on the rock formations, with the result that the rock beneath the bit is crushed and the formation is gradually penetrated. Small amounts of water introduced into the bore hole serve to suspend the cuttings and leave a more or less fresh rock surface exposed for each blow. The bit is frequently withdrawn and a bailer is intermittently introduced to remove the debris which, if allowed to accumulate, would soon render the drilling ineffective. When a water-bearing formation is encountered and the hole is filled with water, the downward velocity of the bit is greatly diminished and drilling becomes relatively slow and inefficient. Such water horizons must, therefore, be cased off, entailing considerable expense for pipe and necessarily a reduction in size of the bit subsequently used. Thus each water horizon encountered requires a separate string of casing for sealing-off and the bore hole rapidly tapers to an excessively small bore.

To meet this and other shortcomings of the "standard rig" the "rotary drill" is used. It consists of a hollow drill stem carrying a suitable cutting head at its extremity. The stem is rotated continuously from the surface, with pressure applied at the bit. By means of pumps a suitable drilling fluid is circulated down the hollow drill stem, through orifices at the bit face, and thence back to the surface through the annular space between the drill stem and well bore. This drilling fluid or mud has three distinct functions, namely:

* An elaboration of this paper has been submitted by Mr. Lawton as a master's thesis, University of Pittsburgh.

(1) It washes the rock face clean and carries cuttings to the surface, where the cuttings are settled out in a pit and the mud fluid is used again.

(2) Water, gas and minor oil-bearing formations encountered during drilling are sealed off by the hydrostatic head of the fluid column, controlled by proper adjustment of the density of the drilling fluid.

(3) The pore-sealing effect on the face of the bore hole results in the prevention of caving of loose and unconsolidated formations. Incidentally, the colloidal nature of the mud-fluid serves as a lubricant for the drill stem and for running casing into the hole. Thus the rotary drilling method allows more or less continuous drilling through one or more water or gas horizons without necessarily requiring casing of the bore hole. Obviously, if due care is not taken, or if improper mud-fluid characteristics obtain, temporary suspension of mud-fluid circulation may cause settling of the cuttings, resulting in hopelessly jamming the drill stem and preventing subsequent retrieval. Likewise improper control may entail loss of one or more of the important functions of the mud-fluid.

It is apparent that the mud-fluid used is an extremely important item in drilling oil wells and it is of interest to mention that the special mud-fluid used in certain drilling operations in the past has occasionally represented an expenditure as high as $30,000; hence, the importance attached to this apparently trivial item of drilling equipment. It is not too extravagant a statment to say that the proper quality and control of the mud-fluid is the key to successful drilling.

It is only recently that petroleum technologists have had a full appreciation of the importance of adequate control by chemico-physical means of rotary drilling fluid in the technique of boring oil wells. Any drilling fluid may be regarded as a system of three main constituents, as follows: water, colloidal matter, consisting of both gel-forming and non-gel-forming colloids, and larger particles suspended in the medium. Of the two types of colloid present, the gel-forming type is definitely the more important, for this group imparts the necessary properties of an efficient drilling fluid, i. e., stability of suspension, viscosity, pore-sealing, and gel-characteristics in general. Colloids of the non-gelling type, together with larger particles, contribute somewhat to viscosity but more particularly to the density of the drilling fluid. Adequate control is a problem of colloid physics and it will be shown how the gel-type of colloid present responds to treatment, giving desirable and necessary characteristics of drilling fluid, modified, of course, by the inert material always present.

In this paper we shall restrict ourselves to clay drilling fluid and a consideration of the problem of lowering the viscosity. It is generally desirable in drilling practice to maintain the viscosity in the neighborhood of 15 centipoises, as measured by the Stormer viscometer at 10 revolutions per second. For a given clay this naturally limits the maximum density of drilling fluid that can be mixed without treatment, but the mud may then lack the desired weight and corrective measures become imperative. A consideration of the basic functions required of a drilling fluid, insofar as cuttings and gas are

concerned, together with the characteristics that are required to fulfill these functions, leads us to the following classification:

1. Carrying cuttings
   Requires alone or in combination:
   - High density, High velocity } in the turbulent region
   - High density, High viscosity } in the viscous region.

2. Elutriation or dropping of cuttings
   Requires alone or in combination:
   - Low density
   - Low viscosity
   - Low velocity
   - Time, or reduction of path.

3. Release of gas
   Requires:
   - Low viscosity
   - Time, or reduction of path
   - Other factors outside the scope of this paper.

From the above it is seen that the requirements for carrying of the cuttings up the bore hole and settling of the cuttings in the mud pit are diametrically opposed; also, one requirement for release of gas is opposed to one of the requirements for carrying of cuttings in the viscous region. Low viscosity favors two of the desirable functions of the mud-fluid, while a high viscosity favors one, the capacity for carrying cuttings in viscous flow. With respect to the carrying of cuttings, however, an increase in the mud-stream velocity or the density or both will compensate for lowering the viscosity. When other consderations, such as ease of handling, reduction of the load on the pumps, reclaiming of mud-fluid by elimination of cuttings and gas, etc., are taken into account, it is clear that every effort should be made to keep the viscosity of a mud-fluid at the lowest possible figure. At the same time the mud weight should be maintained at its initial value. The mud could, of course, be diluted with water, with resultant lowering of viscosity and elimination of cuttings in a mechanical separator; but when the mud is rethickened to the weight required, there is danger that the viscosity will be excessive.

## II. Effect of Lowered Viscosity on Pump Pressure

A consideration of the rates of circulation in general drilling practice shows that the critical velocity for mud-fluids ordinarily used is exceeded in all except the 16 and 20 inch holes and that flow is in the turbulent region in the drill pipe in all instances. In the case of turbulent flow Fanning's formula for flow in a circular pipe may be applied to flow in an annular space, such as exists when drilling, provided the mean hydraulic radius is used in the formula.[1] From Fanning's formula, $p = K\eta^{0.2}v^{1.8}\rho^{0.8}$ for unit size of hole of unit depth, where $p$ is the pressure drop, $K$ is a constant, $\eta$ is viscosity, $v$ is velocity and $\rho$ is density, it is seen that pump pressure will be lowered by only the 0.2 power of the drop in viscosity for the same rate of flow and density of mud-fluid. As mentioned above, however, the velocity or density or both must be increased to maintain the same carrying capacity for cuttings when

[1] Engineering Experiment Station, Univ. of Ill., Bulletin No. 222.

the viscosity is decreased. The carrying capacity $v_c$ will be given by $v_c = v - v_s$, where $v$ is the velocity of flow of the mud-fluid and $v_s$ is the velocity of slip of the cuttings. In turbulent flow the velocity of slip between the mud-fluid and cutting is given by Rittinger's formula (for a spherical particle in water):

$$v_s = 9\left(\frac{d(\rho_1 - \rho_2)}{\rho_2}\right)^{1/2}$$

where $v_s$ = velocity of slip in feet/sec., $d$ = diameter of round cutting in feet, $\rho_1$ = density of cutting, lbs./cu. ft., and $\rho_2$ = density of mud-fluid, lbs./cu. ft. (The constant 9 becomes about 3.5 for flat particles.)

It is reasonable to assume that this expression is valid so long as the mud-fluid flow is in the turbulent region, even though the differential velocity of fluid and cutting is sufficient to give a Reynold's number below the critical value. At the maximum the viscosity probably enters the constant (friction factor) of Rittinger's formula to about the 0.2 power or less for liquids showing anomalous viscosity, such as mud-fluid, approaching a value of zero if the drill hole and cuttings are considered to be very rough, as is actually the case.

From Fanning's formula, $p = K_1\eta^{0.2}v^{1.8}$ for unit size and depth of hole and unit density of mud-fluid. The velocity of flow of the fluid is then given by $v = (p/K_1\eta^{0.2})^{1/1.8}$. The velocity of slip $v_s = K_2$, from Rittinger's formula, assuming that the flow of fluid past the cutting is turbulent and that the power index of the viscosity approaches zero for perfect turbulence. Then

$$v_c = (p/K_1\eta^{0.2})^{1/1.8} - K_2.$$

Solving, $p = K_1\eta^{0.2}(v_c + K_2)^{1.8} = K_3\eta^{0.2}$ for constant carrying capacity.

It is to be concluded, then, that pump pressure varies as $\eta^{0.2}$, approximately, for constant carrying capacity of the mud-fluid for cuttings. The final result of lowering the viscosity of the mud-fluid will be a small drop in pump pressure, while the efficiency of removal of the cuttings remains approximately constant.

## III. Thixotropy of Mud-Fluid

The chief objection to the use of a low viscosity mud in the hole, although partially compensated for by increased mud-stream velocity so far as carrying capacity is concerned, is the fact that rate of settling of cuttings will be greater in case circulation is unavoidably stopped. Rate of settling is independent of the height of the column, but the amount settled around the drill stem depends on the height of the column, the time of settling and the content of cuttings. For a given height of column both of the latter factors may be modified. A chemical reagent may be used that gels the mud-fluid quickly while in a static condition, and settling takes place only during the few seconds elapsing before gel-formation. The viscosity of the gel is so large that settling ceases and even large particles are held in suspension. When the pumps are again started the gel breaks easily with restoration of the original low viscosity of the mud-fluid. The content of cuttings may be reduced by using as high a mud-stream velocity as possible for a given rate of drilling and by

settling as many cuttings as possible in the settling pump. It is clear, then, that a chemical reagent must lower the viscosity as much as possible, and cause the formation of a light gel when the mud-fluid is in a static condition. Moreover, the chemical must cause the release of particles larger than 325-mesh only and must stabilize by deflocculation particles smaller than 325-mesh. In this way muds may be reconditioned and the colloid content restored.

## IV. General Consideration of the Viscosity Factor

### 1. Stokes' law

From a qualitative consideration of the functions of a drilling mud it has been shown that viscosity should be as low as possible for a mud of a given weight. The same conclusions follow for dropping cuttings when Stokes' law is considered.

$$V = \frac{2gr^2(\rho_1 - \rho_2)}{9\eta},$$

where $V$ = velocity of fall of a particle, $g$ = acceleration of gravity, $r$ = radius of falling particles, $\rho_1$ = density of particles, $\rho_2$ = density of medium, and $\eta$ = viscosity. In cases where there is no gel formation, hindered settling does not begin until 30 percent solids are present in the fluid,[2] and therefore Stokes' law of unhindered settling applies. For a given weight mud it is clear from the formula that the only practical way to increase the rate of removal of cuttings is to decrease the viscosity. The same equation with the same conclusion applies in the case of release of entrained gas from a mud, since the rising bubbles of gas also follow Stokes' law approximately.

### 2. Differential settling

In order to reclaim or recondition clay mud-fluids it is necessary to settle out the fine sand as well as large cuttings, while at the same time recovering the finely-dispersed clay particles. For practical purposes we may say that all particles larger than 325-mesh should be settled out while all particles smaller than these should be held in suspension. Suppose that we divide the mud particles into two main sizes, large ones and small ones, designated by $L$ and $S$, respectively, then we shall have for the velocities of fall of each size

$$V_L = \frac{Kr_L^2(\rho_L - \rho_2)}{\eta} \text{ and } V_S = \frac{Kr_S^2(\rho_S - \rho_2)}{\eta}.$$

We are interested in the practical case that the large particles settle out faster, i. e., $V_L > V_S$, and wish to investigate the magnitude of $V_L - V_S$.

$$V_L - V_S = \frac{K}{\eta}\left[r_L^2(\rho_L - \rho_2) - r_S^2(\rho_S - \rho_2)\right]$$

$$= \frac{Kr_L^2}{\eta}\left\{\rho_L - \left(\frac{r_s}{r_L}\right)^2\rho_S - \rho_2\left[1 - \left(\frac{r_s}{r_L}\right)^2\right]\right\} > 0$$

[2] Bignell, Oil and Gas Journal, May 22, 1930, p. 50.

$$\therefore \quad \rho_L > \left(\frac{r_s}{r_L}\right)^2 \rho_S + \rho_2\left[1 - \left(\frac{r_s}{r_L}\right)^2\right].$$

This immediately tells us that we should make $\rho_L$ as large as possible and $\rho_S$, $\rho_2$, and $\eta$ as small as possible. Since $\rho_L$ and $\rho_S$ cannot be subjected to arbitrary variation under practical conditions the only conclusion possible is that the density and viscosity of the mud should be decreased as much as possible in order to obtain a large differential effect. In drilling practice it is important to maintain the density as high as possible in order to seal off gas and water horizons by the hydrostatic column of the mud-fluid. Our only remaining variable, then, is viscosity. If we assume that $r_L \gg r_S$ then $(r_S/r_L)^2$ is negligible and

$$V_L - V_S = \frac{K r_L^2}{\eta}(\rho_L - \rho_2).$$

The final solution again tells us that viscosity must be lowered as much as possible in order to reclaim or recondition a mud-fluid by elimination of the larger particles by settling.

The only remaining factor affecting the release of sand and cuttings is the length of path of the mud-fluid in the time allowed for release by the capacity of the mud storage system. It will be shown that certain chemical reagents lower the viscosity markedly; as a result, cuttings settle rapidly and the length of the settling ditch ordinarily used is more than sufficient. Baffles should be installed in the settling trough in order to break the gel that tends to form and the velocity of the mud-fluid stream should be large enough to counteract any tendency toward an increase in viscosity through gel-formation. The optimum velocity should be just short of turbulence in order to effect the most efficient release of cuttings and possible entrained gas.

## V. Chemical Reagents

The equations given above take no account of electrical or adsorption effects with which colloid physics deals, for Stokes' law does not apply to particles of colloidal size. By taking advantage of colloid phenomena much better results for differential settling are obtained than the above equations would indicate. Certain chemicals that are strongly adsorbed by the colloidal clay particles cause remarkable lowering in viscosity of the mud-fluid. As a result, the large cuttings and sand particles settle out quickly; the small particles are stabilized by deflocculation, a light gel is usually formed within a short time, and the small particles remain in suspension indefinitely.

The authors are not aware of any comprehensive quantitative data published on chemical treatment of mud-fluids. Duckham[3] refers briefly to the use of soda, soda ash and sodium aluminate, with very few data based on measurements. Doherty, Gill, and Parsons[4] mention several chemicals that have been used in field practice. A commercial preparation known as "Sta-

[3] Duckham, Inst. Pet. Tech., March 1931, p. 153.

[4] Doherty, Gill, and Parsons, Oil and Gas Journal, June 11, 1931, p. 52.

bilite" appeared on the market about a year ago. Certain private reports that have come to our attention recommended sodium silicate as the best reagent, but the viscosity was lowered only slightly for the actual mud and it was necessary to thin with water in order to eliminate the sand. The same results could be obtained without chemical by diluting a little further with water.

## VI. Lowering of Viscosity by Chemical Reagents

In our experimental work, apparent viscosities of the treated mud-fluids were measured with a Stormer viscometer, in a manner previously described.[5] All samples were screened through a 70-mesh screen before making viscosity measurements, for larger particles interfere with the free rotation of the drum of the instrument. Practically all the measurements were made with Pierce Junction rather than with Seminole mud-fluid, for the initial viscosity of the former is considerably higher. The weights of these mud-fluids were 9.75#/gal. and 10.31#/gal., respectively, after screening through 70-mesh.

Fig. 1 shows the lowering in viscosity of both Pierce Junction and Seminole mud when treated with caustic soda. It is seen that concentrations (ex-

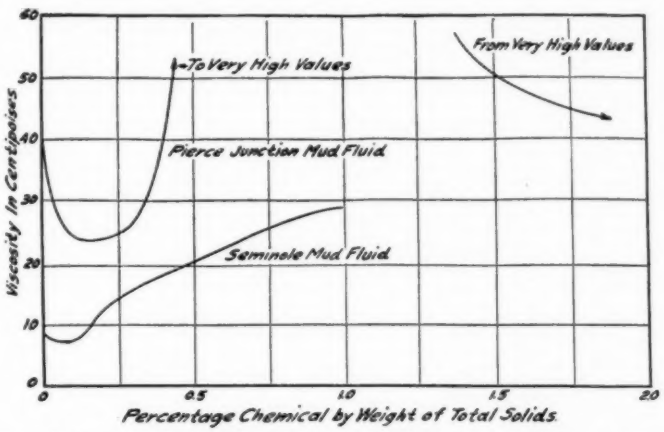



Fig. 1. Clay mud fluids treated with caustic soda. Viscosity in centipoises vs. percentage caustic soda by weight of solids. Temperature 25°C.

pressed in percentage of caustic by weight of total solids of the mud) of caustic soda between 0.1 and 0.3 have a marked lowering effect on the viscosity of Pierce Junction mud, but that the effect is not so large for Seminole mud. Increasing concentrations cause a rapid rise in viscosity and the mud becomes plastic. At still greater concentrations the viscosity drops and the mud is flocculated, settling rapidly as a sludge. We would expect the hydroxide ion to be the most effective of the monovalent ions because of its strong adsorption. Further work must be done to compare with multivalent anions.

Fig. 2 summarizes the data for caustic soda, "Mud-It", (a commercial pro-

[5] Ambrose and Loomis, Physics **1**, 129 (1931).

duct containing sodium aluminate with excess caustic soda), sodium oleate (soap), Stabilite, sap-brown (a commercial product containing humates

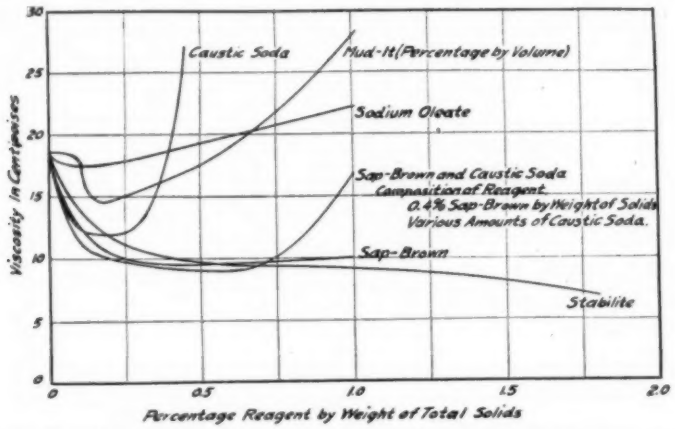



Fig. 2. Pierce Junction mud-fluid treated with various chemical reagents. Viscosity vs. percentage reagent by weight of solids. Temperature 25°C.

and lignates), and sap-brown mixed with caustic soda. It is seen that Mud-It and sodium oleate have little merit in lowering viscosity while sap-brown and Stabilite are superior to caustic soda both with respect to lowering the viscosity and elimination of the rapid increase in viscosity shown by caustic soda with increase in concentration.

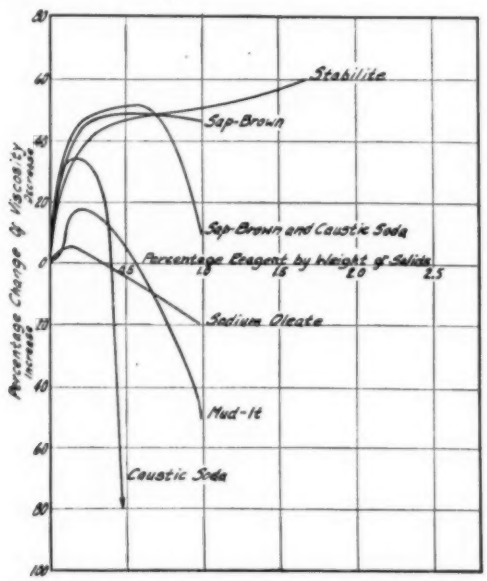



Fig. 3. Pierce Junction mud-fluid treated with various chemical reagents. Percentage change of viscosity vs. percentage reagent by weight of total solids.

Fig. 3 shows the percentage lowering of viscosity as a function of the concentration of chemical.

Fig. 4 presents comparative results obtained on a different sample of Pierce Junction mud with sap-brown, sap-brown plus caustic soda, and Sta-

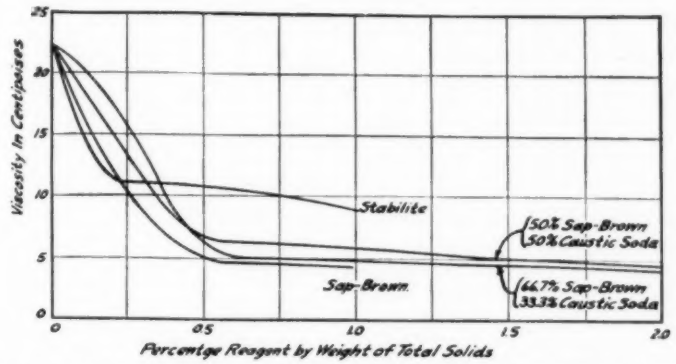



Fig. 4. Pierce Junction mud-fluid treated with various chemical reagents. Viscosity in centipoises vs. percentage reagent by weight of total solids. Temperature 25°C.

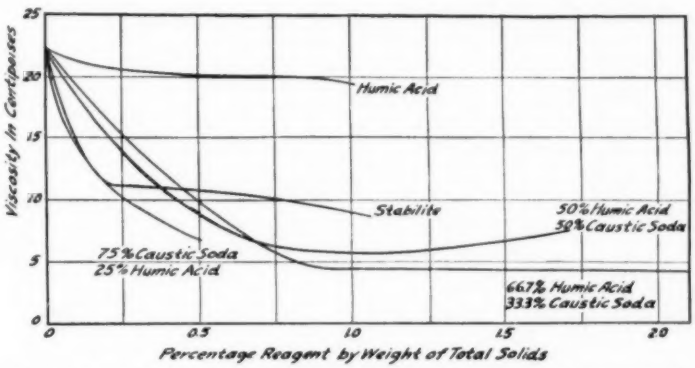



Fig. 5. Pierce Junction mud-fluid treated with various chemical reagents. Viscosity in centipoises vs. percentage reagent by weight of total solids.

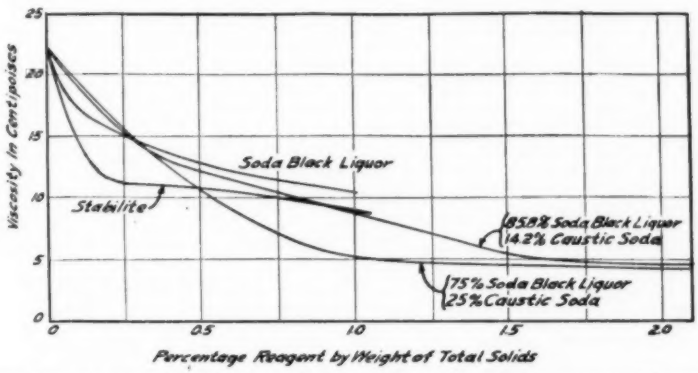



Fig. 6. Pierce Junction mud-fluid treated with various chemical reagents. Viscosity in centipoises vs. percentage reagent by weight of total solids. Temperature 25°C.

bilite. The two samples of mud were taken in the field at different times and showed differences in color and in initial viscosity.

Fig. 5 compares the efficiency of humic acid, humic acid plus varying amounts of caustic soda, and Stabilite.

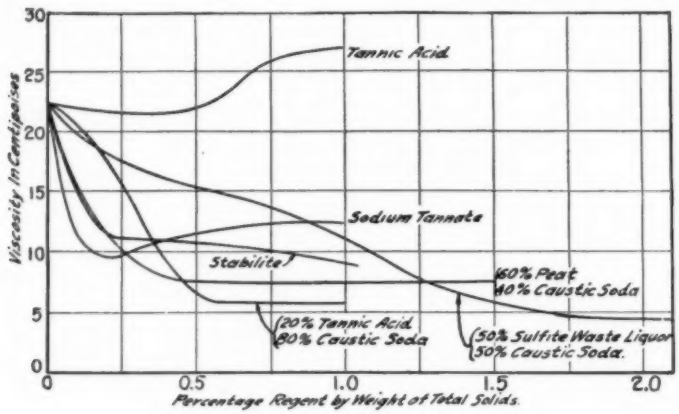



Fig. 7. Pierce Junction mud-fluid treated with various chemical reagents. Viscosity in centipoises vs. percentage reagent by weight of total solids. Temperature 25°C.

Fig. 6 illustrates the action of Stabilite with soda black liquor, both alone and with addition of caustic soda. The soda black liquor is a by-product of the paper industry, and contains mostly lignins and carbohydrates with sodium carbonate.

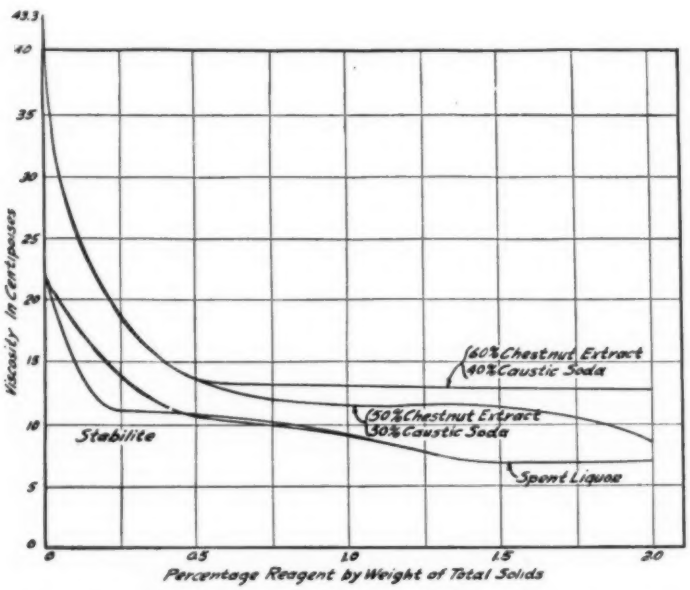



Fig. 8. Pierce Junction mud-fluid treated with various chemical reagents. Viscosity in centipoises vs. percentage reagent by weight of total solids. Temperature 25°C.

Fig. 7 shows that tannic acid alone has only a slight effect in lowering viscosity, while tannic acid with increasing amounts of caustic soda added is

superior to Stabilite. Peat with caustic soda also gives excellent results; peat contains humic acid chiefly. Sulfite liquor, also a by-product of the paper industry, gives excellent results, but somewhat greater concentrations are necessary.

Fig. 8 compares Stabilite with chestnut wood extract, containing tannates and carbohydrates, and spent liquor; the latter contains the material extracted from rice hulls by caustic soda and is a by-product of the Lake Charles Products Company.

From these data it is evident that the organic materials used are much superior to the inorganic reagents for the purpose of lowering the viscosity of clay muds. The degree of gelling the mud is almost entirely dependent on the amount of caustic soda used with a given organic chemical. The mixtures referred to in the figures form light gels with the clay.

It should be emphasized that the water used in these experiments was of very low mineral content. Similar results in field practice cannot be expected if brines are employed.

## VII. Acknowledgments

It is a pleasure to acknowledge the valuable suggestions of Dr. P. D. Foote, Mr. R. J. S. Pigott, and Mr. R. D. Wyckoff of the Gulf Research Laboratory in the preparation of this paper. Our discussion of the flow of mud-fluid is based largely upon the suggestions of Mr. Pigott.

# Radon Content of Soil Gas

By HOLBROOK G. BOTSET, *Gulf Research Laboratory, Pittsburgh, Pa. and*
PAUL WEAVER, *Gulf Production Company, Houston, Texas*

(Received March 17, 1932)

An electroscope and technique for measuring the radon content of soil gas are described. Measurements were made to answer several questions regarding the radon content of soil gas. It was found that there is no systematic increase of radon content with depth of penetration into a soil, that soil formations appear to have more or less characteristic radon contents, that high radon content of a soil gas is not necessarily an indication of the presence of a fault, and vice versa; and finally it is found that there is great local variation of radon content of soil gases.

## INTRODUCTION

IN RECENT years there has been some discussion as to the applicability of radioactive methods in geophysical prospecting. Various claims have been made by different investigators as to the geological significance of the measurement of the radon content of soils. It has been stated for instance that a fault will be evidenced by a higher radon concentration in the overlying soil,[1] and that the radon content of soil in an oil bearing region is higher.[2] In order to test the usefulness of radioactivity measurements for geological purposes an extensive study was made of the radon content of soils over regions whose geology and lithology were rather well known. Measurements were made over faulted areas, salt domes, and areas which were geologically monotonous.

The measurements were made in soils which were developed over sedimentary formations consisting of predominantly sandy clays with some slightly consolidated sands, all of tertiary age, by climates with annual rainfall varying from 31 to 50 inches, and average temperature of about 68 degrees Fahrenheit. Most of the measurements were made in wooded areas, but some were in cultivated fields and some in natural prairies.

In a few cases the measurements were from the bed rock below the soil, but even in these cases the sampling indicated that the formation had been effected by oxidation. With these few exceptions, all measurements are from the "*B*" layer of the soil and are above the level of ground water.

## DESCRIPTION OF APPARATUS AND TECHNIQUE

The instrument used was a modified form of the Ambronn[3] electroscope. The technique of sampling and measuring the soil gas which was developed after a preliminary study was as follows.

[1] F. Müller, Zeits. f. Geoph. **3**, 330–6 (1927).

[2] L. N. Bogoyavlensky, Bull. Inst. Pract. Geoph. No. 3, pp. 113–124. Leningrad, 1927.

[3] Ambronn, Methods of Applied Geophysics. Elements of Geophysics, English translation by Cobb, p. 119.

A pointed steel driving rod about 1–1/8 inches in diameter was driven vertically into the ground to a depth of about two feet. This was then withdrawn and an earth auger 1–1/8 inches in diameter was used to continue the hole to the desired depth. The purpose of the earth auger was to prevent compacting the soil and thus reducing the porosity in the immediate vicinity of the hole. Two depths of hole were used in the surveys: 3 ft. 9 in. and 5 feet. The sampling tube which was then pushed into this hole consisted of a length of 5/8 inch iron pipe containing perforations near the lower end to allow passage of the soil gas. The lower end of the tube was plugged to prevent it from becoming stopped up when being pushed into the hole. Slightly above the middle of this pipe was welded an 8 inch length of 1–1/4 inch iron pipe. This was to provide a seal against the side of the hole when the sampling tube was in position. In order to reinforce this seal a spiral of 1/4 inch square rod was welded to the outside of this larger section. A hose connection was fastened to the top of the sampling tube. The gas sample was thus pumped from near the bottom of the hole, through a rubber tube into the ionization chamber of the electroscope, by means of a hand operated piston pump.

The electroscope consisted of a single fiber electrometer mounted on a cylindrical ionization chamber of about one liter capacity. The lower end of this ionization chamber consisted of a movable piston, thus simplifying the emptying and filling of the electroscope. There were two ports in the ionization chamber, one, an exhaust and the other an intake port. In passing from the soil to the electroscope the gas passed through two $CaCl_2$ drying tubes.

The procedure in sampling and measuring was as follows: The piston of the electroscope was pushed up, thus emptying the ionization chamber. Before connecting the hose from the sampling tube to the ionization chamber a few strokes of the pump sufficed to replace the air in the system with ground air. The hose was then connected to the intake port, the exhaust port closed and air pumped from the hole until the ionization chamber was completely filled. This method of filling by pumping into the ionization chamber assured that the leakage in the system, if any, would be outward, thus preventing the dilution of the sample with atmospheric air, which might very easily occur if the sample were taken by the suction produced by pulling down on the ionization chamber piston. As soon as the electroscope was filled the intake port was closed and a stopwatch snapped. Exactly five minutes later the rate of movement of the fiber across a given portion of the scale was read. Because of the nonlinearity of the scale the same portion was always used in making the readings.

Since in this method, not only the radon but also the active deposit resulting from the decay of radon, was measured, the radioactivity was rapidly changing during the first few minutes after filling the ionization chamber. This change (increase) in activity reaches a maximum after about four hours; however, it reaches about sixty percent of the maximum in the first five minutes. For this reason the reading was made just five minutes after filling the ionization chamber. An additional reason for waiting five minutes before reading the fiber movement was the elimination of the effect of the possible

presence of thoron. Thoron, being the decay product of thorium $X$ is frequently present in soil gas. Since thoron has a half life of only 54 seconds, it will have decayed to about 3 percent of its initial value in five minutes and consequently will have little effect upon the radon measurement. If the reading is made less than five minutes after filling the ionization chamber and thoron is present, its presence and varying concentration during the reading will result in an apparently high radon content which will be considerably in error. A reading of the blank leak was made immediately before each measurement in order to be sure that there was no undue contamination of the electroscope from the preceding samples. Since a measurement was made approximately every fifteen minutes, some contamination was always carried over from the preceding readings. This was evidenced by a gradually increasing blank leakage during the course of the day. The over night rest of the instrument allowed most of the residual active deposit to decay and the blank in the morning was always much improved over that of the previous evening.

The results in this paper are expressed in scale divisions per second of movement of the electrometer fiber $\times 1000$. Calibration showed this to correspond to a radon content of the gas of $15.39 \times 10^{-12}$ curies of radon. To reduce this to a value for the radium content of the soil is difficult because the quantity of radium is very small and some of it is probably inside the grains of the soil, hence it is not certain that all of the radon escapes to the soil gas (according to Satterly, only one sixth to one twentieth of the radon formed in the soil gets into the soil gas). Also differences in porosity of the soil will cause different holes to draw gas from greater or less volumes of soil. In some cases the porosity was found so low that no gas could be extracted with the pump. It is therefore only an approximation which can be made, but it is concluded that 100 on our electrometer scale in this paper will represent about $0.57 \times 10^{-12}$ curies of radon per cc of soil. This electrometer scale which is here called "electroscope reading" is the rate of movement of the electrometer

TABLE I.[4] *Comparison of values for radon content of soil gases.*

| Author | Soil type | Content in $\frac{\text{curies}}{\text{cm}^3}$ |
|---|---|---|
| A. Gockel | | $0.07$ to $0.28 \times 10^{-12}$ |
| J. Satterly | | $0.25 \times 10^{-12}$ |
| J. Satterly | Swamp | $0.3 \times 10^{-12}$ |
| J. C. Anderson | Sandstone | $\left.\begin{matrix}0.24\\1.35\end{matrix}\right\} \times 10^{-6}$ g Th |
| K. Kähler | Sand | $0.008 \times 10^{-12}$ curies. |
| J. Olujie | Loam | $1.20 \times 10^{-12}$ curies. |
| P. Ludewig and E. Lorenzer | Uranium containing rock | $0.3$ to $18.2 \times 10^{-12}$ |

fiber in divisions per second, multiplied by 1000. Consequently an electroscope reading of 100 corresponds to $1539 \times 10^{-12}$ curies of radon in the ionization chamber. The above figure of $0.57 \times 10^{-12}$ curies per cc of soil is obtained by assuming a porosity of 33 percent. If we assume the density of the

[4] "Radioaktivität", St. Meyer and Schweidler, p. 593.

soil grains as 2.6 we obtain $0.33\times10^{-12}$ curies/gm of mineral for an electroscope reading of 100. If we assume that about one sixth of the radon formed gets into the soil gas we obtain a value of $1.98\times10^{-12}$ gm radium per gram of mineral. The average value for the radium content of sedimentary rocks is $1.4\times10^{-12}$ gm/gm.

Fletcher found by analysis, a radium content in two soil samples from Dublin, of 1.3 and $2.9\times10^{-12}$ gm/gm.[5] Table I shows for comparison the values of radon content of soil gases as obtained by different authors.

## Results

Data on the following points were obtained during the survey: (1) The depth at which the radon content is representative of the formation. (2) Comparison of results from areas of different soil texture. (3) Effect of variation in moisture content on readings. (4) Whether or not a given formation has a characteristic radon content. (5) Whether or not an outcropping fault is characterized by a higher radon content. (6) Whether the radon content is correlated with any other property of the soil.

### 1. The depth at which the radon content is characteristic of the formation

A survey was made over an area in which the soil was fairly uniform in texture: clay at the surface with some sand toward the bottom. Two holes were made at each station, a shallow hole, 3 ft. 9 in. and a deep hole, 5 ft. deep. The deep hole was placed about one foot from the shallow hole. The results given in Table II show that for the depths at which samples were taken, there is no correlation with depth.

Table II. *Radon measurements from deep and shallow holes. Electroscope reading (division per second $\times$ 1000).*

| Shallow | Deep | Shallow | Deep |
|---|---|---|---|
| 1 | 0 | 102 | 89 |
| 121 | 131 | 0 | 156 |
| 78 | 66 | 168 | 123 |
| 86 | 58 | 31 | 2 |
| 158 | 131 | 0 | 0 |
| 0 | 0 | | |

It has been stated by other investigators[3] that there is an increase in radon content with depth of penetration into a soil. In none of our surveys did we

Table III. *Radon measurements from deep and shallow holes. Electroscope reading (divisions per second $\times$ 1000).*

| Shallow | Deep | Shallow | Deep |
|---|---|---|---|
| 142 | 150 | 49 | 142 |
| 132 | 0 | 252 | 170 |
| 101 | 57 | 241 | 215 |
| 108 | 252 | 136 | 78 |
| 134 | 166 | 106 | 89 |
| 76 | 87 | 61 | 151 |

[5] A. Gockel, Die Radioaktivität von Boden and Quellen, p. 36.

find evidence confirming this statement. Table III shows the data from another survey in an entirly different region than that shown in Table II.

Again we notice that there is no correlation with depth.

It was noticed that in regions where the soil near the surface was sandy the readings from shallow holes were in general lower than those from deep holes. This was due to one or both of two causes: the sandy nature of the soil may have allowed the gas to be diluted by atmospheric air through soil breathing or the seal at the sampling tube may not have been sufficiently tight to prevent leakage.

In order to determine whether there really was leakage past the seal a method was developed for testing each hole. This consisted in introducing acetylene gas around the shoulder of the sampling tube just below the surface of the ground. The acetylene was prepared from a little calcium carbide and water. A U.S. Bureau of Mines carbon monoxide detector (without the charcoal adsorbent) was applied at the outlet end of the pump. If this detector, capable of detecting a very small fraction of a percent of $C_2H_2$, showed any discoloration the seal was obviously leaking and the hole was abandoned. This very sensitive method of detecting leaks showed that a large percentage of shallow holes in the more sandy type of soil, leaked and consequently gave a lower radon content than the true value. It would thus appear that where the soil is dense or where there is a dense layer at or near the surface the radon content near the surface will be characteristic of the place from which the sample is taken. In more sandy regions the higher porosity of the soil may permit dilution of the sample with atmospheric air. It will be shown later that there is great local variation of radon content and where the deeper gave lower readings than the shallow holes this is probably due to such local variation. In the Reynosa gravel in those places where it was possible to drill a hole the readings were zero. It is impossible to say whether this was due to the composition of the gravel or to leakage through the large pores since at that time the method of testing for leaks had not been developed.

**2. Effect of soil texture on radon content of soil gas**

Measurements were made on soils ranging in texture from coarse gravel to sticky gumbo clay. In coarse sandy soil both high and low readings were obtained, and the same was true for the clays. In those regions where a sandy soil was encountered below a denser more clayey layer the radon content was always quite high. This is probably due to the sealing action of the denser clay layer which prevents diffusion to the atmosphere.[6] However, in general, it may be stated that the texture of the soil has little if anything to do with the radon content of the soil gas.

**3. Effect of variation of moisture on radon content**

A survey was made in January after a protracted rain over an area which was quite low and swampy. In March of the same year after much drier weather the same area was surveyed again. In some cases the sampling tube

[6] E. Blanck, Handbuch der Bodenlehre (Springer, 1930), p. 304, Vol. VI.

was put into a hole of the original survey. The comparison of results from these two surveys is shown in Table IV.

TABLE IV. *Comparison of radon measurements in dry and wet weather.*

| Electroscope reading First survey (div./sec.×1000) | Electroscope reading Second survey (div./sec.×1000) | Electroscope reading First survey (div./sec.×1000) | Electroscope reading Second survey (div./sec.×1000) |
|---|---|---|---|
| 51 | 15 | 406 | 7 |
| 64 | 115 | 195 | 23 |
| 115 | 134 | 195 | 45* |
| 98 | 93 | 195 | 117* |
| 177 | 85 | 208 | 80 |
| 177 | 88 | 28 | 57* |
| 406 | 0 | | |

* These readings taken by putting the sampling tube down the original holes made on the first survey.

It will be noted that with but three exceptions the results of the second survey are all lower than those of the first. This tends to confirm the hypothesis that in times of heavy rainfall the radon is apt to be sealed in the ground and consequently to come more nearly to equilibrium with the contained radium salts.

**4. Does a geological formation have a characteristic radon content?**

A number of profiles was run across several formation contacts. It was noticed in several cases that as the formation contact was approached from

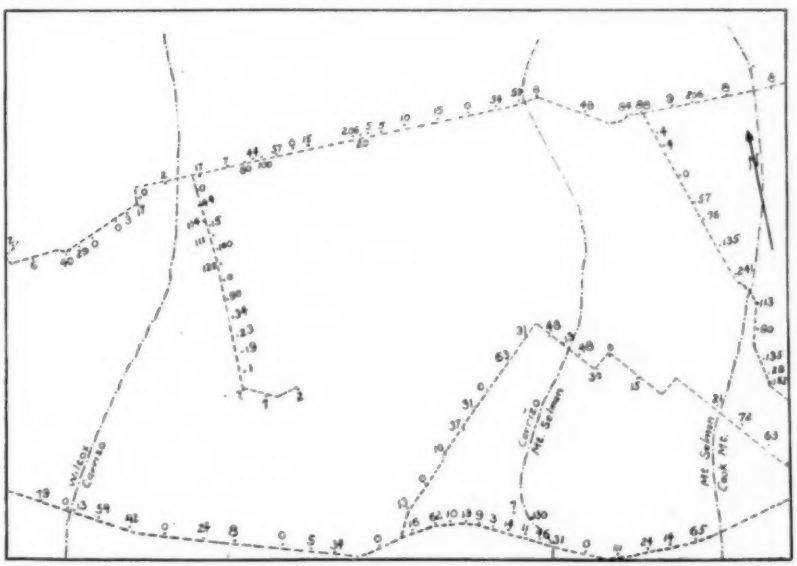



Fig. 1. Radon survey over four soil horizons. Numbers indicate electroscope readings (div. per second × 1000). Scale 1:75,000.

the upper part of the formation the radon content increased. The formations over which the survey was made are the Wilcox, Carrizo, Mt. Selman and

Cook Mountain. The formation contacts were those between the formations in the order named above. The accompanying map, Fig. 1, shows the radon readings across these formations and contacts.

The averages of the readings in each formation are as follows:

| | | | |
|---|---|---|---|
| Wilcox | 34 | Mt. Selman | 56 |
| Carrizo | 30 | Cook Mountain | 82 |

It would seem from this that although individual readings varied between very wide limits the average over a formation may have some significance. The dip of the beds was to the east southeast; the lower part of the surface of each bed is toward the west in Fig. 1. It will be seen that the radon content of the soil near the bottom of the bed seems in general to be higher than that in the upper part. This is especially noticeable on another survey made across the contact between the Wilcox and the Midway (Fig. 2).

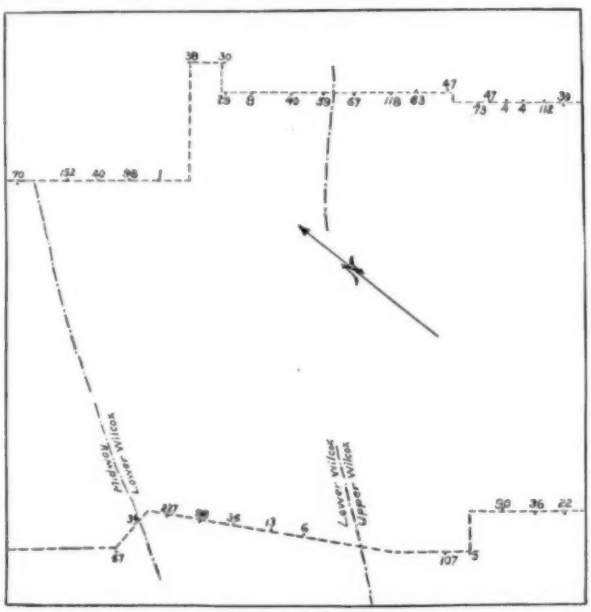



Fig. 2. Radon survey over three soil horizons. Numbers indicate electroscope readings (div. per second × 1000). Scale 1:63,100.

There was less variation in the Cook Mountain than in any other formation. The Carrizo, a non-marine formation showed the lowest radon content of any horizon studied. It is composed of coarse sand, generally very loose. It contains cross bedded sandstones at intervals and high readings were found near and in these sandstones. The Wilcox in general is quite variable as far as radon content is concerned. Rapid changes have been noted by the geologist in the texture of this formation, and tests of surface water and surface specific resistivity have confirmed this variability.

In general the results of the survey indicate that formations may have characteristic radon contents. Additional surveys over other formations would be desirable to obtain further data on this point.

**5. Does an increase in radon content indicate the presence of a fault?**

It is stated by Ambronn and others that certain "Graben" and faults in Europe are detectable by the large increase in radon content of the gas in the soil above them. In order to test this statement, fairly extensive surveys were made over areas whose subsurface geology was quite well known. Several parallel profiles were run across a fault whose surface trace was known to within a few feet. These are shown in Fig. 3. Although three of the profiles do show a high reading at the true fault line, $A$, the data would also indicate a fault at $B$. However, it is known that there is no fault at $B$; consequently it is seen that the data are very misleading.

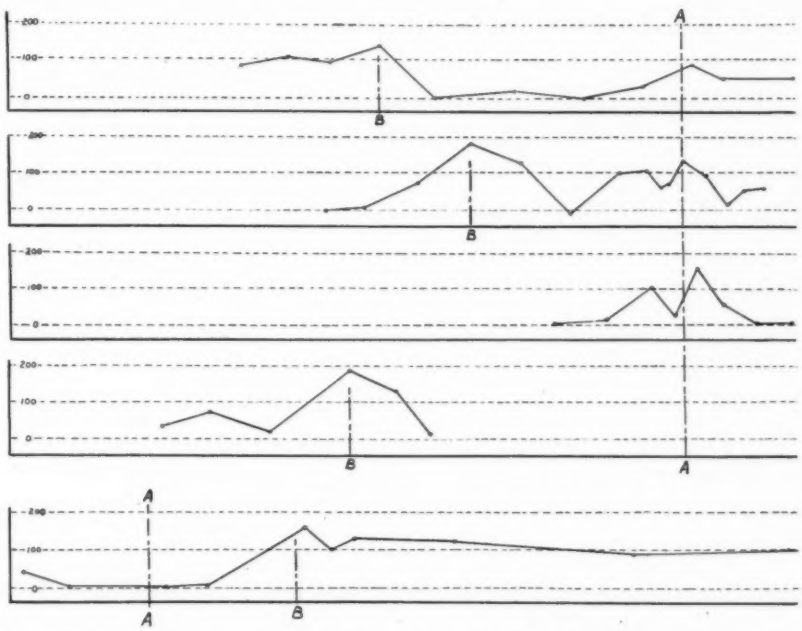



Fig. 3. Radon surveys across a known fault. Vertical scale, electroscope readings. Horizontal scale $1''=1400'$. $A-A$, position of known fault. $B$, erroneous fault which would be indicated by radon survey.

Additional measurements were made at the outcrop of a very large fault. Measurements made in the soil twenty-five and fifty feet from this fault gave readings of 30 and 122 respectively. A sample of gas was taken from the very fault plane itself and gave a reading of 99. It is evident that there is no enrichment in radon resulting from circulation up the fault plane in this case. In regions of Europe where similar measurements have been made the outcropping faults were indicated by higher radon concentrations. This is due no doubt to the fact that upward circulation of water along those fault planes has carried radon and other radioactive substances to the surface and produced the observed effects. It seems safe to say, however, that unless there is upward circulation of water to the surface through a fault plane, there will be no increased radon content in the superposed soil.

In order to study further the extent to which the radon content of a soil is the result of extremely local conditions a survey consisting of 16 deep and 12 shallow holes was made over an area 30×35 feet. As seen in Fig. 4, the holes were arranged in the form of two grids, one superposed on the other and so placed that the holes of one grid were equidistant from the holes of the other. The soil here seemed to be quite uniform; the surface was entirely covered by sod. In drilling the holes a light sandy clay of a bright reddish color was first encountered. The soil at the bottom of the holes was lighter in

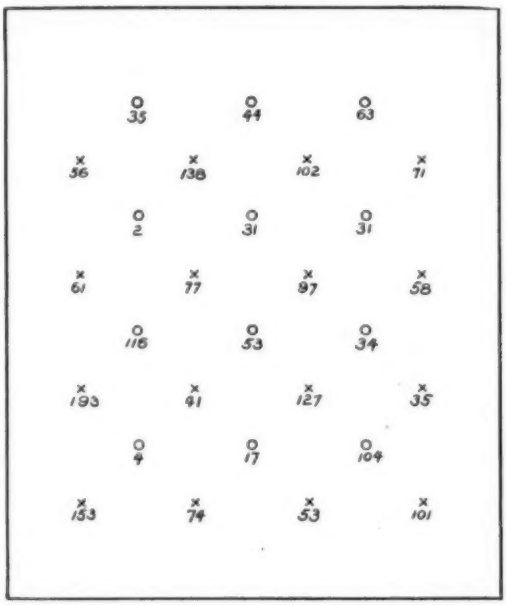



Fig. 4. Radon content of soil over limited area. Figures indicate electroscope readings. × = holes 5 feet deep ○ = holes 3 feet 9 inches deep. Outside dimensions of rectangle 30′×35′.

color, varying between a reddish and a yellowish orange. The bottoms of the holes were very sandy.

As will be seen from the figure the variation in radon content among these holes was as great as was encountered anywhere else. In other words, ordinary local variations in soil will be sufficient to mask any other variations of greater geological significance. In attempting to determine a characteristic radon value for a given formation a great many measurements must be made in order to average out any local effects.

**6. Is radon content correlated with any other property of the soil?**

The measurement of the radon content of the soil is really the measurement of something quite different. It is a measurement indirectly of the uranium content of the soil. Radon is a decay product of radium which in turn is a decay product of uranium. Radon has only a very brief life, any quantity being practically completely decayed in forty days. Consequently it must be constantly renewed in the soil gas. It is renewed by the decay of radium

which is a solid element and of much longer life, since a given quantity of radium will be almost entirely decayed in 16,000 years. The soil at any appreciable depth has probably been more or less undisturbed for much longer periods than that so the ultimate source of the radon is the primary parent element uranium which has a life of billions of years.

The uranium and radium being solids are contained in the grains of the soil and consequently only a comparatively small fraction of the total radon produced is able to escape from the grains and get into the soil air. Hence the radon content of a soil gas will depend not only on the parent radioactive material contained in the soil grains, but also upon the size and nature of the surface of the grains. Since the radon is indirectly a measure of the uranium present it appears that there should be some correlation, other conditions remaining constant, between the radon and certain heavy mineral constituents of the soil, such as zircons and biotites.

## Conclusions

A study of the radon content of soil gases over extended areas indicates that there is no correlation between radon content and depth, that the radon content of a given formation may have a characteristic value, although considerably more data on this point would be desirable. It is also indicated that abnormally high radon content of soil gas does not necessarily show the presence of a fault, and that there are very great local variations in radon content of such a nature as to suggest that there may be some correlation between radon content and certain heavy mineral constituents of the soil.

The authors wish to acknowledge the help and cooperation of Dr. Paul D. Foote, Dr. A. E. Ruark, and Mr. R. L. Chenault in carrying out this investigation.

## Bibliography

St. Meyer and E. Schweidler, Radioaktivität.
L. B. Smyth, Phil. Mag. (6) **24,** 632 (1912).
J. R. Wright and O. F. Smith, Phys. Rev. (2), **5,** 459 (1915).
V. F. Hess and W. Schmidt, Phys. Zeits. **19,** 109 (1918).
P. Ludewig and E. Lorenzer, Zeits. f. Physik **22,** 178 (1924).
A. Gockel, Phys. Zeits. **9,** 304 (1908).
A. Gockel, Die Radioaktivität von Boden u. Quellen, Vieweg, Braunschweig, 1914.
Ambronn-Cobb, Elements of Geophysics.
F. Müller, Zeits. f. Geoph. **3,** 330–6 (1927).
E. Blanck, Handbuch der Bodenlehre, Vol. VI, Springer, 1930, p. 304.

# Speed, Speed Factor and Power Input of Different Designs of Diffusion Pumps, and Remarks on Measurements of Speed

By T. L. Ho*
*Norman Bridge Laboratory of Physics, California Institute of Technology*

(Received March 5, 1932)

The speed factor of a diffusion pump is defined as the ratio of its real speed to the ideal maximum speed. The fact that the speed factor is much less than unity is due to three causes, namely, the vapor dispersed through the slit, the nozzle chamber resistance and the diffuse reflection of gas molecules at the boundary of the vapor-stream and at the condensing wall. Some effects on the speed are considered as being independent of the configuration. The following general conclusions result from the measurement of the speeds of different configurations. (a) In order to reduce the amount of dispersed vapor by the slit, the vapor nozzle should be divergent instead of being parallel; (b) in order to reduce the chamber resistance the space surrounding the slit should be large and the nozzle chamber should be cold; (c) there exists for each design a certain slit area at which the speed is maximum, but there has been no general rule to predetermine it; (d) in designing a large pump either a narrow throat annular nozzle or multiple nozzles can be used. In measuring speeds some precautions should not be overlooked, unless the results will be falsified: Either the determined speed will be too small or too large (pseudospeed).

## I. Introduction

ALL present constructions of diffusion pumps can be classified as one single or a combination of the three following general types. *A*, *B or C*, in Fig. 1, show a longitudinal section of the principal part of each type which characterizes its inverted and titled forms. One cannot say generally that one type is more efficient than the other and that pumps of the same type are equally efficient. Each type can have different designs, hence the

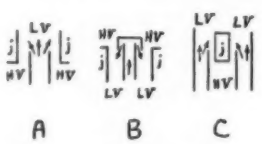



Fig. 1. The principal parts of three types of diffusion pump. The arrows denote the directions of the vapor, *LV* the low vacuum, *HV* the high vacuum and *j* the cooling system.

efficiency is different for different individual pumps. Very few previous investigators tried to explain the reasons of the efficiency of their designs. However, the need of the high speed pumps in modern research renders a discussion of the physical principles involved in the design of efficient pumps a necessity. The present investigation is to determine experimentally the speed, speed factor and power input of different designs and to draw the general conclusions by discussing the experimental results.

* Research Fellow of the China Foundation for the Promotion of Education and Culture.

## II. Definition of Speed Factor

The volume of a gas escaping per unit time, through a slit of area $A$ on a thin wall of a gas container, into a perfect vacuum or into a space in which the escaping gas molecules are completely and continually washed away by some arrangement, is approximately:

$$V = \left(\frac{RT}{2\pi M}\right)^{1/2} \cdot A \tag{1}$$

where $R$ is the gas constant, $T$ the absolute temperature of the gas and $M$ the molecular weight of the gas, obtained by considering the number of gas molecules striking a surface $A$ per unit time. This is also the ideal maximum speed of a diffusion pump having a slit of area $A$. For air at room temperature, $V \cong 11.7$ liters per cm$^2$. We define the speed factor as the ratio of the actual speed of a diffusion pump to the ideal maximum speed. This speed factor is of the utmost importance in pump design.

## III. General Considerations

Considering a gas passing through a slit into a vapor stream as shown in Fig. 2, in the space above the slit, (corresponding to the nozzle chamber) according to its different states of motion the vapor is to be considered as consisting of two types: directional vapor and nondirectional vapor. The directional vapor is directly radiated in from the slit and has a resultant up-

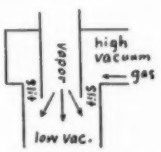



Fig. 2. The principal part of a diffusion pump of type $A$ inverted.

ward momentum, while the nondirectional vapor is produced by re-evaporation and collision and possesses thermal agitation which depends on the temperature of the chamber. We first consider that the gas molecules diffuse through the directional vapor alone. Gaede[1] and more recently Molthan[2] treated this kind of diffusion process analytically. They reached the conclusion that the number of gas molecules striking the slit area $A$ per unit time is

$$z = \tfrac{1}{4}\alpha N \Omega A \tag{2}$$

where $N$ and $\Omega$ are respectively the concentration and the mean velocity of the gas molecules; and $\alpha = f(\lambda/d) < 1$, where $\lambda$ = mean free path of gas molecules in vapor and $d$ = width of slit, which increases with $\lambda/d$. However, $N$ cannot be the same throughout the space above the slit whenever there is a flow of gas. So, we regard this $N$ only as the molecular concentration of the gas near the slit. Different molecular concentrations along the path of the

[1] Gaede, Ann. d. Physik. **46,** 357 (1915).
[2] Molthan, Zeits. f. Tech. Physik. **7,** 377 (1926).

gas in the chamber can be considered as due to the resultant resistance offered by the chamber and the vapor in it, mostly the nondirectional vapor. Before discussing this resistance we have to derive another coefficient.

If there were no chances to return for gas molecules entering the slit, there would be $z$ gas molecules carried away per unit time by the vapor stream. Actually, there are some gas molecules which rebound from the vapor molecules and also from the condensing wall below the slit.

Let $n$ be the number of the gas molecules returning through the slit per unit time, and assume this is directly proportional to $z$, i.e., $n=kz$. Then the total number of the gas molecules extracted per unit time is

$$\begin{aligned} Z &= z - n \\ &= \tfrac{1}{4}\alpha\gamma A N\Omega, \end{aligned} \tag{4}$$

where $\gamma=1-k,<1$, and depends largely on the shape of the nozzle and the intensity distribution of the vapor stream below the slit.

To express the speed of the pump, substituting in (4) $N=\rho/m$ and $\Omega=(8p/\pi\rho)^{1/2}$, where $m$, $p$, $\rho$ are the mass of a molecule, the pressure, and the density of the gas respectively, and using the perfect gas law, we obtain the volume of the gas extracted per unit time,

$$S = \alpha\gamma A\left(\frac{RT}{2\pi M}\right)^{1/2}. \tag{5}$$

This is the speed at a point near the slit.

Now we come again to the chamber resistance. There is a speed difference between the inlet mouth and the region near the slit due to the chamber resistance.

Let $S_r$ be the real speed (i.e., the speed at the inlet of the pump) and $R_c$ the chamber resistance. Then, by the definition:[3]

$$\frac{1}{S_r} = \frac{1}{S} + R_c,$$

or

$$S_r = \frac{S}{1+SR_c}. \tag{6}$$

But $R_c=1/U$, $U$ being the conductance of the chamber. Regarding a rarefied gas passing through a vapor as a gas permeating a porous substance, $U$ can be written as:

$$U=(RT/M)^{1/2}KF \text{ (dimension of chamber, density of vapor)},$$

where $K$ is a numerical constant, and $F$ is increasing with increasing cross-section of the chamber and with decreasing vapor density.

Then we have, by substituting the values of $S$ and $R_c$ in (6),

$$S_r = \alpha\beta\gamma A\left(\frac{RT}{2\pi M}\right)^{1/2}, \tag{7}$$

[3] Dunoyer, La Technique Du Vide, p. 132 (1924).

where

$$\beta = \frac{1}{1 + \alpha\gamma A \dfrac{1}{(2\pi)^{1/2}KF}},$$

which increases with $F$.

By comparing (1) and (7), the product of the coefficients $\alpha\beta\gamma$ is the speed factor of a pump of given area $A$. Hence in designing a pump of large speed one should make each of the coefficients $\alpha$, $\beta$, $\gamma$ as large as possible.

In the above expression of speed, the assumption is made that the intensity of the vapor stream is so large and the fore pressure is so low that there are only very few gas molecules diffusing back from the fore vacuum to the slit.

## IV. General Experimental Results

Some effects on the speed, which are independent of the design of the pump, have been experimentally studied by the author, of which only a brief summary shall be given.

(a) The speed is independent of the fore pressure when the latter is low, whereas the upper limit of the fore pressure is quite critical. The same result was shown by Langmuir[4] and can be seen also in the speed-pressure graphs obtained by Hickman and Sanford.[5] This critical fore pressure increases with the vapor pressure in the boiler.

(b) If the fore pressure is kept below its critical value, there exists an optimum value for the vapor pressure, above which the speed decreases again, as stated first by Gaede. This is due to the fact that $\alpha$ and $\beta$ increase with decreasing vapor pressure and the weak vapor stream causes an increase of gas molecules reflected on it and a back diffusion from the fore vacuum side.

(c) At a given optimum range of vapor pressure in the boiler the speed is nearly inversely proportional to the square root of the molecular weight of the gas to be extracted. This shows, from (7), that the speed factor is nearly the same for different gases. Indeed, for a given width of the slit, Gaede's expression of $\alpha$ is a function of the mean free path, $\lambda$, of the gas molecules in the vapor, which depends largely on the diameter and velocity of the gas molecules. However, when $\lambda$ is larger than the width of the slit, $\alpha$ is almost independent of $\lambda$ and it is the same for different gases.

(d) The speed is independent of the pressure of the gas being extracted. This is to be expected, since (7) does not involve $p$. Of course, the range of the pressure in which the speed is constant is limited by the boundary conditions of (7). Gaede also experimentally verified it by his observations on the first diffusion pump.

(e) Varying the temperature of the cooling system, the speed first increases with decreasing temperature and then reaches constancy, since the

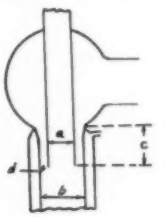
[4] Langmuir, Gen. Elec. Rev. **19**, 1060 (1916).
[5] Hickman and Sanford, R.S.I. **1**, 140 (1930).

coefficient $\beta$ is obviously increased by lowering the temperature of the chamber, and also the rapid condensation of the vapor will increase the coefficients $\alpha$ and $\gamma$.

(f) The speed of the pump filled with *n*-butyl phthalate is higher than that of the same pump filled with mercury. For the *n*-butyl phthalate has a lower vapor pressure than mercury as has been determined by Hickman and Sanford,[6] hence it causes an increase of $\alpha$ and $\beta$.

## V. Experimental Results of Different Designs

The word "design" applies only to the size and shape of the principal part (i.e., the vapor nozzle and its surroundings) of the pump, the boiler, the heater and the vapor transmission tube being disregarded.




| | a | b | c | d |
|---|---|---|---|---|
| Fig. 3. Pump 1A | 1.7 | 2.4 | 2.5 | 0.25 cm |
| Pump 1B | 3.0 | 3.5 | 1.4 | 0.15 |
| Pump 1C | 3.5 | 4.3 | 1.4 | 0.25 |

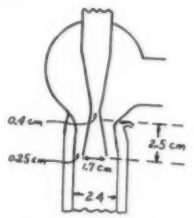



Fig. 4. Pump 2.

The pumps which have been tested are listed with their measured data in Table I. Some of them are schematically shown in the diagrams (Fig. 3 to Fig. 13). Pump 7 (Fig. 9) was constructed in this laboratory by imitating

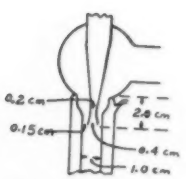



Fig. 5. Pump 3.

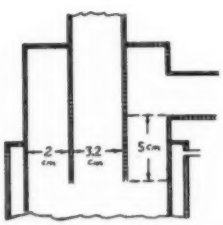



Fig. 6. Pump 4 (Steel).

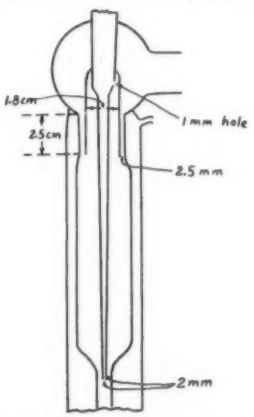



Fig. 7. Pump 5 (Two stages).

Gaede and Keesom's[7] high speed pump. Pump 8 (Fig. 10) is Hickman and Sanford's design but with a bigger nozzle chamber and with a removable condenser. Pumps 10, 11, 12 and 13 are new designs of this laboratory.[8]

[6] Hickman and Sanford, J. Phys. Chem. **34,** 637 (1930).
[7] Gaede and Keesom, Commun. Leiden **18,** No. 195a, (1928).
[8] Ho, R.S.I. **3,** 133 (1932).

For the low speed pumps, the speeds were measured by a flowmeter, which is calibrated for air and other gases, and constantly checked by the

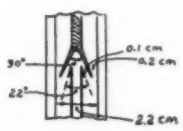



Fig. 8. Pump 6 (Steel).

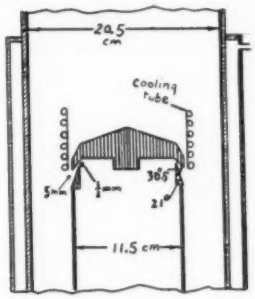



Fig. 9. Pump 7 (Gaede and Keesom's high speed pump).

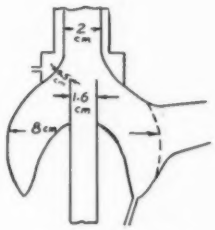
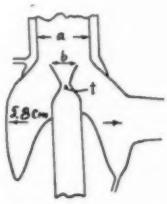



Fig. 10. Pump 8.




| | a | b | t |
|---|---|---|---|
| Fig. 11. Pump $9_A$ | 2.0 | 1.0 | 0.3 cm |
| Pump $9_B$ | 3.6 | 2.2 | 0.6 cm |

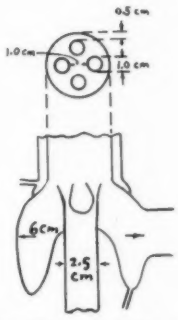



Fig. 12. Pump 10 (4 nozzles).

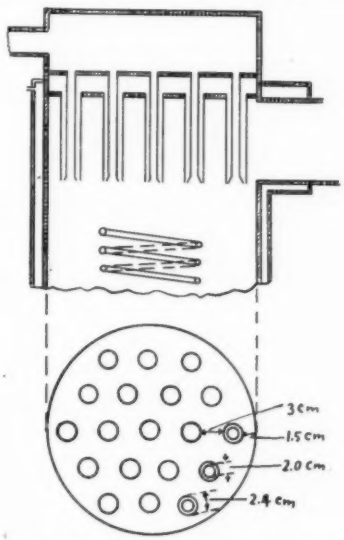



Fig. 13. Pump 13 (Lauritsen's 19 nozzle pump).

constant volume method. For the high speed pumps, the speeds were measured by the mercury-pellet method and constantly checked by comparison with the known speed of a fore pump.

The results in Table I all refer to mercury as a fuel. The speeds were specified as (1) at the position along the inlet tube and close to the nozzle chamber, (2) constant at pressures from $10^{-3}$ to $10^{-5}$mm Hg, (3) against a fore pressure of below 0.1 mm Hg, (4) the water cooling was at room temperature

TABLE I.

| Pump | Form | Slit area | Watts input | Speed (for air) | Speed factor |
|---|---|---|---|---|---|
| $1_A$ | Straight nozzle Small chamber | 1.54 cm² | 115 | 1.01 L | 0.056 |
| $1_B$ | " | 1.48 | 140 | 3.5 | 0.20 |
| $1_C$ | " | 2.9 | 140 | 5.7 | 0.17 |
| 2 | Divergent nozzle Small chamber | 1.54 | 115 | 1.57 | 0.086 |
| 3 | " | 0.27 | 110 | 0.57 | 0.18 |
| 4 | Straight nozzle Steel | 30.0 | 200 | 6.6 | 0.019 |
| 5 | 2 stages Small chamber | 1.56 | 200 | 0.82 | 0.045 |
| 6 | Small Steel | 1.26 | 300 | 1.16 | 0.079 |
| 7 | Large Steel | 227.0 | 2200 | 105 | 0.040 |
| 8 | Small Upward | 2.5 | 60 | 3.2 | 0.11 |
| $9_A$ | " | 2.4 | 60 | 4.0 | 0.14 |
| $9_B$ | Large Upward | 6.4 | 140 | 14.5 | 0.20 |
| 10 | 4 nozzles Upward | 9.5 | 140 | 28.0 | 0.26 |
| 11 | 4 nozzles | 9.5 | 140 | 33.0 | 0.30 |
| 12 | 7 nozzles | 22.8 | 200 | 55 | 0.21 |
| 13 | 19 nozzles | 155 | 700 | 320 | 0.18 |

and the temperature difference between the inlet and the outlet water was about 2°C. The slit area of each single nozzle pump was calculated by taking the shortest distance from the edge of the nozzle to the condensing wall. The watts input includes the very small heat losses by radiation and conduction.

## VI. Discussions and Conclusions

### A. Divergent nozzle

Although the parallel vapor beam cannot possibly be obtained in Crawford's pumps[9] with divergent nozzle as pointed out by Gaede,[10] the vapor on passing through the divergent nozzle is expanding at the throat and being partially rectified by the wall of the nozzle, thus the average upward component of the momentum of the vapor molecules after leaving the nozzle is greatly reduced. Consequently, the relative amount of the vapor dispersed through the slit into the chamber is diminished. This, combined with the effective downward vapor stream, will make the speed factor greater. This is why the speeds of pumps $1_A$ and 8 are respectively higher than those of the pumps 2 and $9_A$, and why the divergent jet is adopted in the design of a high speed pump.

[9] Crawford, Phys. Rev. **10**, 557 (1917).
[10] Gaede, Zeits. f. tech. Physik **4**, 348 (1923).

**B. Chamber resistance**

The chamber resistance which affects mainly the coefficient $\beta$ of the speed factor consists of two factors, as mentioned in III, one due to the dimensions of the chamber and the other due to the vapor in the chamber. In order to minimize this resistance it is necessary to enlarge the space near the slit and to keep the chamber cold. The speed is low if the nozzle is thrust deeply into the condensing tube as in the pumps $1_A$, 2 and 5; while the designs of pumps 8 and 9 are of low chamber resistance; and their speeds would be still higher if they were of "vertically downward form", because in the "vertically upward form" the condensed mercury and the boiler underneath heat the chamber considerably.

**C. Slit area**

From the data we can see that the speed factor increases with the increasing of the width of the slit, if the other parts of the design do not change much. This is due partly to the increasing of the coefficient $\gamma$ because the number of the gas molecules reflected back through each unit area per unit time is decreased, and partly to the increasing of the coefficient $\alpha$, since $\lambda/\alpha$ is decreased. Gaede explained this by the term of "extracting power" (Aufnahmefähigkeit) of the vapor stream. Of course, there must be given a maximum value to the product $\alpha\gamma A$ by varying $A$. This is shown partly by pumps $1_B$, $1_C$ and 4 which are all of straight nozzle. The slit of pump 4 might be too large and that of pump $1_B$ too small, since their speeds are not as high as they should be by comparing their sizes with pump $1_C$.

**D. Large pumps**

To obtain a high speed we have to enlarge the size of the pump. But the high speed cannot be obtained by enlarging all the parts of a small pump in proportion. Obviously, the width of the slit cannot be enlarged much if the other parts are enlarged in proportion. So the slit area of a large pump is usually small in comparison with its size (i.e., the area of the cross-section of the condensing tube.) There have been two ways to enlarge the efficient area. (i) By using a convergent nozzle with a very narrow throat (Fig. 9) and applying a high temperature vapor, Gaede and Keesom were able to get an effective vapor stream to operate under a large slit area. (ii) By using several small nozzles in place of a single large one and leaving sufficient space between the nozzles for the additional path of the gas (Fig. 13), an efficient enlarged slit area is obtained. The large speed factor and low power input of the multiple nozzle pumps are remarkable.

## VII. Remarks on Measurements of Speed

Whatever method is used to measure the speed of a diffusion pump the pressure of the gas under extraction must be known. In order to get the actual speed directly, the pressure should be measured at the inlet mouth of the pump. If the pressure gauge is connected at any other point than along the inlet mouth a very serious departure from the true speed will sometimes occur.

**(a) Under-value speed**

The speed of a pump can be obtained by measuring the apparent speed at the far end of a connecting tube of known dimensions and applying resistance[3] equation to calculate the real speed. This method is obviously limited by the condition that the connecting tube should be so large that the ideal maximum speed, defined by Eq. (1), through an area equal to the area of cross section of the tube is much higher than the speed of the pump. Otherwise the measurement can never reach the real speed of the pump. The author once overlooked the above condition and got a speed of 11 liters instead of 105 liters per second of the pump 7, using a connecting tube of only 1.6 cm in diameter.

**(b) Over-value speed**

The departure due to the over-correction is also considerable, if the pressure gauge is directly connected to the nozzle chamber. Gaede stated, for his three-stage steel pump, that at a point adjacent to the slit a speed of 60 liters per second was obtained, but at the wide mouth of the inlet tube the added resistance (i.e., chamber resistance) reduced the result to 15 liters per second. From his statement it is seen that 15 liters per second is the real speed of his pump; and 60 liters per second is the over-value speed, but it would have been the true speed if the nozzle chamber and the inlet tube had been enlarged and cooled.

**(c) "Pseudospeed"**

As in above, if the pressure gauge is directly connected to the nozzle chamber of a pump, the inlet of which is not along the same axis as the chamber, the pressure thus measured is not on the main path of the gas and will render the value of the speed meaningless. For, not only can the pressure not be the same throughout the nozzle chamber because of the chamber resistance, but also the pressure distribution cannot be symmetrical with respect to the axis of the chamber (or of the nozzle), especially for the pump of

TABLE II.

| Pump | Speed (for air) | | | Pseudo-speed (for air) | | |
|---|---|---|---|---|---|---|
| | Water cooling | | Still air cooling (copper wire projection) | Water cooling | | Still air cooling (copper wire projection) |
| | With cooling around chamber | Without cooling around chamber | | With cooling around chamber | Without cooling around chamber | |
| 8 | 3.5 L | 2.8 | 0.17 | 3.8 L | 3.3 | 0.50 |
| $9_B$ | 14.5 | 10.0 | | 25.0 | 28.0 | |
| 10 | 28.0 | 15.0 | | 64 | 70 | |

"vertically upward form" with a small and hot chamber. The pseudospeeds of pumps 8, $9_B$ and 10, measured by connecting the pressure gauge at a point opposite to the inlet tube, are shown in Table II. It is interesting to note that the pseudospeeds of pumps $9_B$ and 10 are several times larger than their corresponding true speeds and decreases with the decreasing of the temperature

of the chamber. This is because the hot vapor raised from the lower part of the chamber prevents the gas passing from the inlet tube to the opposite side (where the gauge is connected) so that the partial pressure of the gas (which is measured by the gauge) on that side is low. It can of course be increased by cooling the chamber. The pseudospeed of pump 8, because of the large chamber and its great distance (8 cm) from the boiler, is nearly the true speed when the pump is properly cooled.

Recently, Hickman and Sanford[5] measured the speeds of their pumps by connecting the McLeod gauge to the chamber at a point far from the inlet mouth of the pump (seen from their schematical diagram of the apparatus and their pictures of the pumps). Some of their speed-pressure (fore) curves (for instance, graph 1 and graph 6 in their paper) show that the speed of a pump by "still air cooling" is equal to or higher than that by "fan cooling", the other conditions being the same. This is inexplicable and in contradiction to Crawford's experimental results,[9] and also to one of the general experimental results (i.e., (e) of section IV). Further, as the author tested, for the mercury filled pump the speed is very low if it is only cooled by the still air (with copper wire projections). The speeds Hickman and Sanford measured are probably pseudospeeds which are higher than the corresponding true speeds.

The pseudospeeds shown in Table II are, as true speeds, independent of the pressure of the extracted gas. This shows that the pressure distribution inside the nozzle chamber is the same for different pressures.

In conclusion, the author wishes to express his thanks to Professor A. Goetz and Professor W. R. Smythe for their valuable criticisms, and to Professor C. C. Lauritsen for his helpful suggestions and for the use of his pumps.

# An A.C. Wheatstone Bridge for Audio and Radio-Frequency Measurements

*By* A. HEMINGWAY *and* J. F. MCCLENDON
*Laboratory of Physiological Chemistry, University of Minnesota, Minneapolis*

(Received March 12, 1932)

IN MANY biological problems it is necessary to measure the conductivity and capacity of suspensions, solutions, and tissues with varying frequencies of the alternating current. The method of electrical conductivity and capacity is perhaps the best method of investigating the phenomenon of polarization, which, according to Nernst[1] is the fundamental principle of electrical excitation of tissues. This method has been used by Gildemeister,[2] Hozawa,[3] Lullies,[4] and others. According to Osterhout[5] and Cowdry,[6] the electrical conductivity method is the best quantitative method of following permeability changes of tissues. This is the method employed by McClendon,[7] Brooks,[8] Osterhout.[9] Fricke,[10] and McClendon[11,12] have used the method of electrical conductivity measurements to study the physical structure of blood cells and Fricke[13] and Stewart[14] have used this method to measure the volume of the disperse phase in biological suspensions. In a quantitative investigation of heat dosage in diathermy Hemingway measured the radio-frequency resistance of the human body[15] and living tissues.[16]

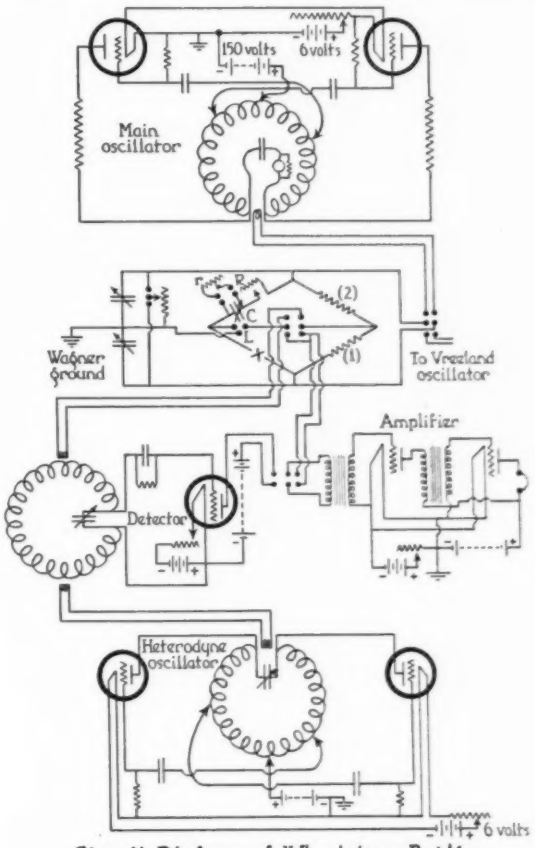
These few examples will illustrate the various biological uses of a Wheatstone bridge, which will be suitable for the measurement of resistance and capacity at various frequencies. The building of a suitable bridge presents considerable difficulty if measurements are to be made at frequencies higher than those of the audio-range. Philippson,[17,18] and Cole[19] have measured by

[1] W. Nernst, Pfluger's Arch. **122,** 275 (1908).
[2] M. Gildemeister, Pfluger's Arch. **219,** 89 (1928).
[3] S. Hozawa, Jour. of Biophysics **1,** 185 (1925).
[4] Lullies, Pfluger's Arch. **221,** 298 (1928).
[5] W. J. V. Osterhout, Jour. Biol. Chem. **36,** 485 (1918).
[6] E. V. Cowdry, Cytology, University of Chicago Press (1924).
[7] J. F. McClendon, Am. Jour. Physiol. **29,** 302 (1912).
[8] S. C. Brooks, Jour. Gen. Physiol. **7,** 349 (1925).
[9] W. J. V. Osterhout, Injury, Recovery and Death in Relation to Conductivity and Permeability, Philadelphia (1922).
[10] H. Fricke, Phys. Rev. **26,** 682 (1925).
[11] J. F. McClendon, Protoplasma **3,** 7 (1927).
[12] J. F. McClendon, Protoplasma **3,** 71 (1927).
[13] H. Fricke, Jour. Gen. Physiol. **6,** 741 (1924).
[14] G. N. Stewart, Am. Jour. Physiol. **90,** 194 (1929).
[15] A. Hemingway, Radiology **14,** 99 (1930).
[16] A. Hemingway and D. C. Collins, Am. Jour. Physiol. **94,** 338 (1932).
[17] M. Philippson, Compt. rend. Soc. Biol. **83,** 1399 (1920).
[18] M. Philippson, Arch. int. Physiol. **18,** 161 (1921).
[19] K. S. Cole, Jour. Gen. Physiol. **12,** 29 (1928).

resonance methods the impedance at various frequencies. The bridge method, however, enables the experimenter to separate the resistance factor of the impedance from the reactance factor, which is of importance in biological work. In the past few years[12] we have been building and perfecting a Wheatstone bridge suitable for the type of measurement mentioned and it is this bridge which is described below.

## The Bridge Circuit

The circuit used is similar to that devised by Stratton,[20] and utilized by Remington.[21] It is shown in Fig. 1. The four arms of the bridge contain: 1 and




Circuit Diagram of Wheatstone Bridge

Fig. 1.

2, equal ratio arms; $X$, the unknown whose resistance and capacity is desired; and $C$, $R$, $r$, a balancing arrangement as shown of two variable decade resistors and a decade capacity. The bridge is also designed to allow an inductance to be placed in series with the unknown in arm $X$, an arrangement used by many investigators.

[20] J. A. Stratton, Jour. Opt. Soc. Am. **13**, 471 (1926).
[21] R. E. Remington, Protoplasma **5**, 339 (1928).

For measurements in the audio-frequency range a Vreeland oscillator is used as a source of alternating current. The characteristics of the Vreeland have been investigated by Taylor and Acree,[22] Morgan and Lammert[23] and others and found to be satisfactory for a. c. measurements due to the minimum of alternating current harmonics. When the Vreeland oscillator is connected to the bridge, an audio-frequency amplifier with ear phones is used in the bridge circuit for balancing.

For radio-frequency measurements a heterodyne method is used. The main oscillator and the heterodyne oscillator produce alternating currents whose frequencies differ by approximately 1000 cycles per second, and these two frequencies are impressed on the detector. The plate current of the detector passes through the primary of the first audio-frequency transformer of the audio-frequency amplifier, which amplifies the beat note of 1000 cycles per second.

Hitherto for our work it has been sufficient to make measurements of resistance and capacity throughout the audio-range of frequencies, and to make a radio-frequency measurement at a single radio-frequency of one million cycles per second. It is our intention, however, in the future to install the proper inductances and capacities in the oscillators and detector circuit for measurements at intermediate frequencies.

### The Oscillators, Detector and Amplifier

For oscillators the two tube Eccles-Jordan type recommended by Stratton[20] has been found to be very satisfactory. The main oscillator has two UX112A tubes with 150 volts on the plates while the heterodyne oscillator has 50 volts on the plates of two UV201A tubes. The detector has a UV201A tube with a plate voltage of 25 volts. An important and necessary feature of these circuits is the inductances which are wound in a toroidal form to reduce the external field. The condenser of the main oscillating circuit is fixed, while those in the detector and heterodyne oscillator are variable for tuning. In the oscillating circuit of the main oscillator is placed a Weston thermocouple milliameter which indicates whether oscillation is taking place and whose current value indicates whether or not the batteries need charging. The amplifier is of the usual type with two stages of amplification. UV201A tubes are used with 50 volts on the plates and transformer coupling.

### Resistors, Condensers, and Ratio Arms

Two decade resistors are used for balancing, one for radio-frequency and the other for audio-frequency, and are connected as shown in Fig. 1. For radio-frequency measurements the 1, 10, and 100 ohm decades are wire woven coils, with a modified Ayrton-Perry winding and constructed especially for our purpose by members of the Radio Engineering Department of the Universtiy of Minnesota. The 1000 and 10,000 ohm units are "El Menco" carbon coated cylindrical resistors. These are calibrated with 1000 cycle alternating

[22] W. A. Taylor and S. F. Acree, Jour. Am. Chem. Soc. **38**, 2396 (1916).
[23] L. R. Morgan and O. M. Lammert, Jour. Am. Chem. Soc. **48**, 1220 (1926).

current when necessary and are kept at fairly constant temperature. For the lower frequency measurements of the audio-range a General Radio type 102 decade resistor is used. The condensers used are Leeds and Northrup precision decade condensers with a variable air condenser for capacities less than 0.001 microfarad.

Ratio arms consist of two "El Menco" units of 100, 1000, or 10,000 ohms for high frequency or two wire resistors wound on long cards with the modified Ayrton-Perry method for audio-frequencies. The two coils are connected in series with three contacts, one from each end of the pair and one from the center, which project downwards into cups of mercury. This method allows of easy interchange of the arms and convenient reversal. The use of equal fixed ratio arms eliminates the difficulties due to drum wound slide wires, mentioned by Jones and Joseph,[24] Morgan and Lammert,[23] Schlesinger and Reed,[25] and Parker.[26]

## Shielding

Proper shielding of the bridge for accurate measurements is essential, especially with radio-frequency currents. This matter has occupied the attention of engineers and physicists in the past few years and several methods are now in general use. One is the older method of Campbell[27] recently adopted by Shackelton,[28] Ferguson[29] and Shedlovsky,[30] and is particularly useful in transmission line work. In this method the individual bridge units are inclosed in a "nest" fashion and the ground currents brought to one or more points of the bridge. This method requires a considerable number of metallic shielding devices which are unnecessary for the type of bridge described in this paper. In the bridge which we have designed the ground leakage has been reduced to a minimum by using spacious metallic containers in which the parts of the bridge are separated from the sheet metal ground containers with a wide air space. This decreases the capacitative short circuiting of a resistor by a metallic conductor in close proximity to it. Galvanized sheet iron is used for shielding. This was chosen, rather than copper which was used temporarily, since measurements were to be taken over a wide range of frequencies and the slight disadvantage due to the lower resistivity of the sheet iron would be compensated by the ferromagnetic properties, which would be effective at the lower frequencies.[31]

## The Wagner Ground

Since Wagner[32] first proposed his now famous earthing device, by which the leak to ground is made equal on both sides of the bridge, several types

[24] G. Jones and R. C. Joseph, Jour. Am. Chem. Soc. **50,** 1049 (1929).
[25] H. I. Schlesinger and F. H. Reed, Jour. Am. Chem. Soc. **41,** 1727 (1919).
[26] H. C. Parker, Jour. Am. Chem. Soc. **45,** 1366 (1923).
[27] G. A. Campbell, Elec. World **43,** 647 (1904).
[28] W. J. Shackelton, Trans. A.I.E.E. **45,** 1266 (1927).
[29] J. E. Ferguson, Trans. A.I.E.E. **48,** 517 (1929).
[30] T. Shedlovsky, Jour. Am. Chem. Soc. **52,** 1793 (1930).
[31] J. H. Morecroft and A. Turner, Jour. Inst. Rad. Eng. **13,** 477 (1925).
[32] K. W. Wagner, Elektrotech. Zeits. **32,** 1001 (1911).

have been used, among which are those of Jones and Joseph,[24] Schlesinger and Reed[25] and Shedlovsky.[30] The one described here is that used by Stratton[22] and consists of two air condensers connected to each of the input leads and to ground, with a variable decade resistance of from 1000 to 110,000 ohms, which can be switched to either input terminal by a d. p. d. t. switch. This arrangement can be adjusted very easily and has proved to be very satisfactory. Only by a careful setting of the Wagner ground can consistent results be obtained with radio-frequency alternating current.

## The Bridge Assembly

The exterior of the assembled bridge is shown in Fig. 2. There are two large sections joined by a cylindrical tube. Within the smaller section is the main high-frequency oscillator and its batteries. Within the cylindrical tube joining the two sections of the bridge are the two leads from the oscillator to

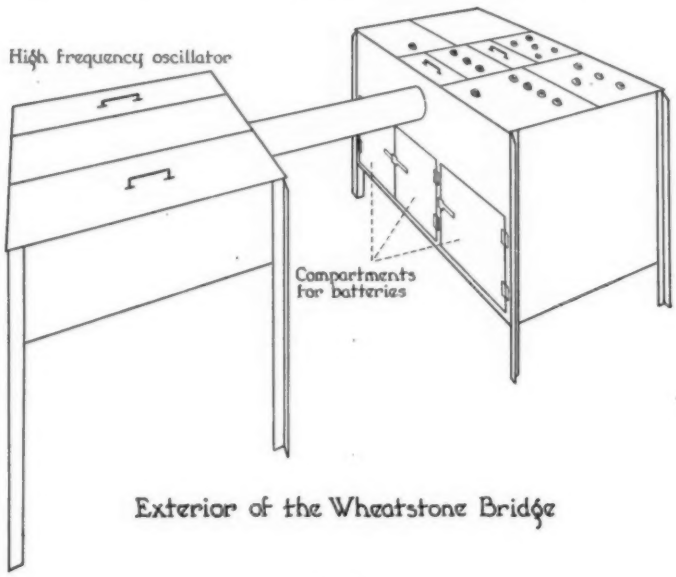



Fig. 2.

the bridge proper. The Vreeland oscillator which furnishes the audio-frequency current is not shown.

The larger section shown in Fig. 2 has two stories. The upper contains the four arms of the bridge, that is the resistors, capacities, the Wagner ground and the constant temperature air bath with the thermostat. The lower story is divided into six compartments, three of which are on the side facing the main oscillator and these three compartments contain batteries, each a 6 volt A battery and B batteries. On the opposite side from the main oscillator of this larger section and in the lower story are three compartments containing the heterodyne oscillator, the detector and the audio-frequency amplifier. Plans showing the arrangements of these two stories are shown in Figs. 3 and 4 and 5. All the apparatus of the upper story is supported on glass rods whose

ends rest on small metallic brackets connected to the sheet iron shield. Between the compartments the wires pass through holes two inches in diameter.

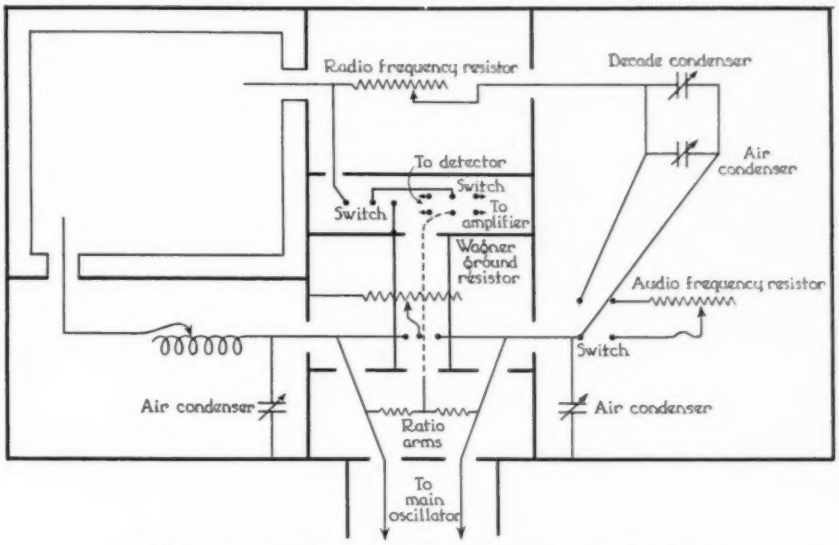



Diagram of Upper Story of Wheatstone Bridge

Fig. 3.

Ebonite or glass rods project from the sliding contacts of the decade dials upwards through small holes in the top of the bridge.

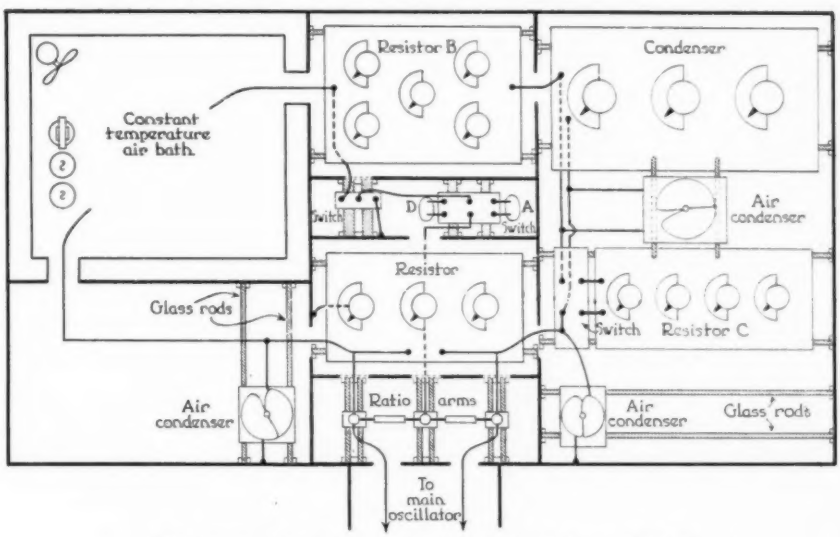



Diagram of Upper Story of Wheatstone Bridge

Fig. 4.

Figs. 3 and 4 show the arrangement of the upper story. Fig. 3 is the schematic wiring diagram and Fig. 4 the actual arrangement. Fig. 4 is shown as

viewed from above with the roof removed. *A* is the low frequency connection; the switch in this position leads to the amplifier directly. In the position *D* the connection is to the detector. Both the amplifier and detector are in the lower story as in Fig. 4. A detailed diagram of the arrangement of the main oscillator, which is the smaller of the two sections, is not given.

## Accuracy

At radio-frequency the accuracy is about 1 percent. To test the bridge manganin wire standards are used. These consist of a single strand of thin manganin wire which runs back and forth between pegs set on opposite edges of a large hardwood board about 1 inch apart. Wound in this way the distributed capacity and inductance are a minimum and the skin effect is small

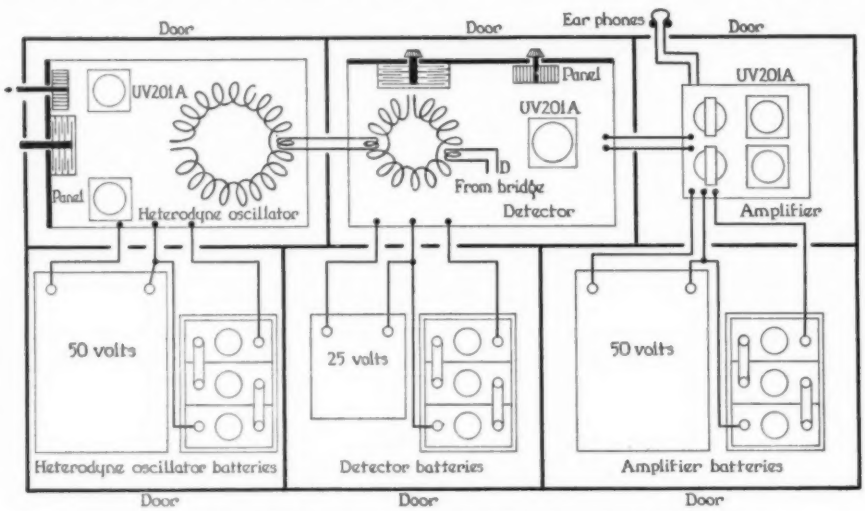



Diagram of Lower Story of Wheatstone Bridge

Fig. 5.

due to the relatively high resistivity of the wire and its small diameter. With these standards the value of the radio-frequency resistance equals that of the 1000 cycle or direct current resistance to within 1 percent. The audio-frequency measurements are accurate to 0.1 percent.

We wish to take this occasion to thank Professor A. Zeleny of the Department of Physics and Professor H. E. Hartig of the Department of Electrical Engineering of the University of Minnesota for the use of apparatus for collaboration purposes and for their asistance and cooperation in this work.

# EDITOR'S COLUMN

## Free Rotation in Molecules

THE rotation of molecular groups (radicals) within complex molecules, about the bonds which hold them together is a phenomenon to which considerable interest is attached. In forthcoming issue of the *Physical Review* a paper[1] by H. Nielson gives an interesting discussion of the quantum theory of an effect of this kind. In a previous paper[2] in the same journal, L. Pauling has considered the theory for an analogous case involving the vibration and rotation of molecules in a crystal lattice. Nielson takes as a model a body dynamically symmetrical about an axis of figure but divided into two parts by a plane perpendicular to the figure axis. The two parts thus formed are assumed to be capable of rotating with respect to each other about the figure axis but subject to a restoring torque tending to keep them in an equilibrium position. A rough mechanical model would be two identical right cylinders placed end to end and joined by a wire passed through them along their common axis, supposing that the wire could twist but not bend, the wire playing the rôle of the chemical bond. There is one important difference between this mechanical model and the actual molecule, namely that in the former there is but one definite equilibrium position being that in which the wire is not twisted at all, for which the restoring torque vanishes, but in the latter there may be two or more equilibrium positions depending upon the symmetry of the arrangement of the atoms within the molecule. In order to take this into account Nielson supposes that the restoring torque can be obtained from a potential of the form $V(\phi')\sim(1-\cos\ m\phi')$ where $\phi'$ is the angle of relative rotation of the two parts and $m$ is an integer. In the range $0\leqq\phi'\leqq 2\pi$ this function vanishes for $2m$ values of $\phi'$, there being $m$ such values in the range $0\leqq\phi'<\pi$ and $m$ further values symmetrically placed in the range $\pi\leqq\phi'<2\pi$.

When Schrödinger's equation is written down for such a system the part depending on the angle $\phi'$ can easily be separated out and is found to be the standard equation for the Mathieu functions familiar in mathematical analysis. If one restricts $\phi'$ to have values only in the neighborhood of one of the equilibrium positions then the equation reduces to that for the harmonic oscillator, in agreement with the fact that for small oscillations the vibrations of a torsion pendulum are simple harmonic. The complete solution of the equation presents difficulties arising from the periodic character of the potential function which is reflected in the periodic character of the solutions of the Mathieu equation. The positions of the energy levels and the transition probabilities are also discussed by the author.

E. HILL

[1] H. Nielson, Phys. Rev. **40,** 445 (1932).
[2] L. Pauling, Phys. Rev. **36,** 460 (1932).

## The Neutron Again

IN VIEW of the importance and interest in the subject it seems suitable to amplify a little the few remarks made in these columns last month on the subject of the neutron.[1] In particular we must mention the important experiments of Mme. I. Curie and M. F. Joliot which form one of the strongest links in the present chain of evidence pointing towards the neutron.[2,3] Since the appearance of Chadwick's note[4] they have presented further evi-

[1] Editor's Column, Physics **2,** 310 (1932).

[2] I. Curie, Comptes Rendus **193,** 1412 (1931); F. Joliot, Comptes Rendus **193,** 1415 (1931).

[3] I. Curie and F. Joliot, Comptes Rendus **194,** 273 and 708 (1932).

[4] J. Chadwick, Nature Feb. 27 (1932) p. 312.

dence of a very convincing nature, convincing at least of the peculiarity of this new radiation.[5] They have made a very careful study of the interesting way in which this radiation is absorbed by substances containing protons or other light nuclei such as helium. The radiation to which we refer is that produced from beryllium atoms when they are bombarded by the $\alpha$-rays from polonium. The anomalous manner in which it is absorbed on passing through paraffin, cellophane, etc. even though its absorption coefficient in lead is very small makes it quite clear that this radiation is not of an electromagnetic nature. Both theory and experiment agree that the ability of a charged particle to take up energy from electromagnetic radiation varies in a strongly inverse manner with the mass of the particle. Thus in the absorption of x-rays by matter, the effect is produced almost entirely by the electrons within the atoms, the nuclei themselves producing only a small effect. But in the present case the reverse is true, and one proton is found to be as efficient as four electrons in absorbing the radiation.

In a recent note evidence of a visual nature has been presented which is quite appealing in its directness.[5] A large Wilson cloud-track chamber is placed in the direct path of the radiation, and within it is placed a mica screen covered with paraffin on which the radiation falls after being filtered through 2 cm of lead. When the chamber is operated tracks are observed due to protons knocked out of the screen by the radiation, these tracks emanating in a bundle from the screen. But in addition to these a number of other tracks have been observed which do not start from the screen but from some point within the body of the chamber itself, that is, from a point within the "water vapor" region of the chamber. These tracks appear to be due to nuclei or protons which have experienced a collision with another particle which itself does an inappreciable amount of ionizing and so leaves no track in the cloud chamber. Photographs of such tracks have been made and cuts of them are shown in the publication. This all agrees very well, at least qualitatively, with the interpretation of the radiation as being composed of neutrons. Being uncharged a neutron would have a very restricted field of force, and a rather close approach would be required for a collision with an atom to be effective in producing ionization. The neutron might travel quite a distance between ionizing collisions.

Similar experiments have been successfully carried out by P. Auger[6] who has studied the paths followed by the ejected protons in a magnetic field. He estimates that the protons have energies from $5\times10^4$ up to several million electron-volts.

In the previous note[1] we neglected to mention one important bit of evidence adduced by Chadwick[4] and Webster;[7] i.e., the experimental observation that the rays emitted from the bombarded beryllium atoms in the same direction as the incident $\alpha$-particles travel is harder (more penetrating) than the radiation emitted in the reverse direction. This is easily understandable on the hypothesis that the radiation is composed of neutrons behaving like particles of unit atomic weight, as the conservation laws for energy and linear momentum show that the neutrons projected in the forward direction by the $\alpha$-particles would have about 50 percent more energy than those ejected the backward direction. On the other hand, one might think that even if the radiation were made up of photons a similar result might hold since Chadwick's computations show that the photon would have an energy of around $14\times10^6$ electron-volts, and such a photon would have a mass equal to about 28 times the rest mass of an electron. As it moves with the velocity of light it carries away an appreciable fraction of the momentum of the incident $\alpha$-particle. If the momentum of the photon varied with its direction of projection, then its frequency would do likewise, and the radiation emitted in different directions with respect to the direction of the incident $\alpha$-particles would have different absorption coefficients. However a direct computation of the difference in frequency for the forward and backward directions of emission of the photon shows that it is negligible under the conditions of the experiments. The evidence thus favors the neutron hypothesis.

E. HILL

[5] I. Curie and F. Joliot, Comptes Rendus **194,** 876 (1932).

[6] P. Auger, Comptes Rendus **194,** 877 (1932).

[7] H. C. Webster, Nature, March 12 (1932) p. 402.

## Cosmic Rays: What Are They?

A CHARGED electroscope placed in air will not keep its charge no matter how carefully the instrument is insulated. The rate of "leak" is small but too large to be simply due to the movement of the charge along the surface of the insulator. The fact that the rate of loss of charge does not depend on the dimensions of the insulator and other tests prove conclusively that the air surrounding the charged body is ionized. These air ions then discharge the electroscope. A part of the ionization in air is produced by the relatively small amount of radioactive material in the air, the underlying soil and rocks and in the material of the electroscope itself. But after all known sources of ionization are taken into account a residual ionization remains. "Cosmic rays are by definition, the cause of that residue of the ionization."[1]

Physicists in various parts of the world have been carrying on experiments to determine the nature of the ionizing agent that produces this residual ionization. It was very soon discovered that the ionizing agent whatever its nature, had an extremely high penetrating power. Regener[2] reports detectable ionization 186 meters and 230 meters below water. The most penetrating of the known gamma-rays spontaneously emitted from radioactive materials could hardly be detected through 20 cm of lead. If one assumes that the ionizing agent observed by Regener is a short electromagnetic wave (or corpuscle, as you please) whose absorption depends on the mass through which it travels, 200 cm of lead would be required to reduce the intensity to the minimum value which he observed.

To account for this penetrating power, Millikan and his school have in the past assumed that the cosmic rays were radiations similar in nature to gamma-rays except much harder, that is, of much higher frequency. Evidence supporting this hypothesis is the constancy of the radiation at all parts of the earth's surface. Millikan and his co-workers have made measurements in California, in the Bolivian High Andes near the equator and at Churchill, Manitoba, 730 miles from the North magnetic pole. If the rays were electrons their distribution at the earth's surface would be modified by the earth's magnetic field. To test further the constancy of the incoming radiation over various parts of the earth, Dr. A. H. Compton and associates are now taking readings at widely separated points, both at high and low altitudes and at high and low latitudes in the northern and southern hemispheres.

We have tacitly assumed that the rays come from outer space into the earth. Observations with an electroscope on the surface of the earth, however, do not determine the direction of travel of the rays. If the source of the rays is on the earth's surface one would expect the rate at which ions are produced in an electroscope to be decreased as one ascended upwards, leaving an absorbing blanket of air between the instrument and the earth. When balloons carrying instruments are sent up 6 to 10 miles it is found that the ionization increases. At sea levels between one and two ions are on the average produced per cc per sec. in air under standard conditions. Piccard, in his famous balloon trip found about 200 per cc per sec. in air (at N.P.T.) at an altitude of 10 miles. Such evidence with that taken below the surface of lakes, rules out the possibility of a terrestrial source for the rays.

Millikan interprets the manner in which the ionization varies with height to be evidence in support of the photon nature of the radiation. The ionization produced when photons pass through a gas is mainly produced by the secondary electrons which are ejected from the gas molecules. At very high frequencies of the photons, theory and experiments indicate that a large fraction of the electrons ejected from the gas atoms will be "Compton-effect" electrons and many of them will follow very closely the original direction of the photon beam. These electrons will have a definite range in a gas, and until one descends below the "top of the atmosphere" to a distance at which the number of the secondary electrons are removed from the path or its immediate neighborhood is equal to the number of them that are produced, the ionization per unit path length will increase. Below such a point the total ionization produced would be decreased since the number of primary photons would be reduced. At the altitude at which the

[1] Darrow, Bell System Tech. Jour. Jan. 1932, gives a complete review and discussion of this subject, covering the literature to a late date in 1931.

[2] Regener, Die Naturwissenschaften **15,** 183 (1929). Zeits. f. Physik **74,** 433 (1932).

beam of photons produces a maximum ionization in a closed vessel, the number of ionizing sources removed from the beam is equal to the number produced in it. At that point, the beam is said to be in equilibrium with its secondaries. A variation with height in a manner consistent with the photon hypothesis has been found in the balloon experiments. The number of ions produced increased to a maximum at a given height and then decreased at higher altitudes.

With the assurance that the cosmic rays were electromagnetic waves, Millikan has used various absorption coefficients to get an explanation of his curves giving the variation of ionization with depth below water. To get a calculated curve to fit the empirical curve, he assumes three or more groups of waves having different absorption coefficients. By means of formulae relating the absorption coefficient to frequency, the frequencies of various cosmic rays are computed and by the use of $E=h\nu$, their energies are given. To find a source for the energy of the magnitude ($2\times10^8$ electron volts) given by these calculations presented some difficulty. Millikan suggested that "in the depths of space" the celestially common elements, helium, oxygen and silicon were created by the union of the proper number of protons and electrons. In this process some mass is lost as is indicated by Aston's curves of atomic masses. If one assumes that the disappearance of a mass $M$ is accompanied by the appearance of radiation whose energy is given by $MC^2=E$ where $E=h\nu$ is the energy and $C=3\times10^{10}$ cm/sec., a correlation between the energies of the cosmic rays and the suggested atom-building processes is obtained. This suggestion has led to numerous speculative discussions concerning the "cycle of the Universe". It allows one to assume the mass of the stars to be continually changing into radiation without necessarily assuming that the final state of the Universe is one of radiation existing alone. Since the cosmic rays do not vary with the position of the observer with respect to the Milky Way or any other celestial body, it is assumed that the atom-building processes take place in free space away from the stars. And the suggestion is made that possibly a process the reverse of that taking place in the stars, is going on in the depths of space, that is radiant energy condensing into atoms.[3]

Regener[3] concludes that the most penetrating of the radiation he detected arises from the annihilation of an entire He atom and the creation of electromagnetic radiation. The wave-length of such radiation is about three millionths of an angstrom unit.

The above conclusions are all based on the results obtained in closed vessels filled with gases at atmospheric or higher pressures. The total ionization in the vessel is determined and no information given concerning the details of the production of this ionization. By use of the Wilson cloud expansion apparatus it is possible to detect visually the path of individual ionizing agents, in particular alpha particles, swift electrons and protons. Skobelzyn several years ago observed paths in an expansion chamber in which gamma-rays were liberating secondary electrons. A magnetic field was applied to bend the electrons in spirals and from the radius of curvature of the spiral, the speed of the electrons could be measured. A magnetic field was applied to bend the electrons in spirals and from the radius of curvature of the spiral, the speed of the electrons could be measured. The gamma rays were those from known radioactive materials. Some of the paths were far too straight to be accounted for by electrons produced by the gamma rays. This observation led to similar experiments to study the cosmic rays, since it was assumed that these straight paths were caused by the cosmic rays themselves or by secondary rays which in turn were produced by the cosmic rays.

Mott-Smith and Locher[4] combined a Wilson cloud expansion apparatus with a Geiger-Müller counter. Bothe and Kolhörster[5] had shown that a Geiger-Müller counter would respond to individual ionizing particles (whatever their nature). These physicists had placed one counter above another and obtained numerous simultaneous discharges in them. Mott-Smith and Locher desired to learn whether or not the agents causing the simultaneous discharges in the counters were the same as those producing the visible paths in the apparatus of Bothe and Kolhörster. They placed the Wilson expansion chamber below

[3] Millikan and Cameron, Phys. Rev. **32,** 533 (1928).

[4] Mott-Smith and Locher, Phys. Rev. **38,** 1399 (1931).

[5] Bothe and Kolhörster, Zeit. f. Physik **56,** 751 (1929).

the Geiger-Müller counters and concluded that visible tracks were produced by the same agents that produced the simultaneous discharges in the Geiger-Müller counters. Such visible straight line tracks have previously been detected only when produced by material particles; protons, alpha-particles or electrons.

Does this then mean that cosmic rays are charged material particles? If so magnetic fields of very high strength may be used to measure their velocity and energy. Curtiss earlier had obtained results which indicated that the number of coincidences of discharges in two Geiger-Müller counters was reduced when a transverse magnetic field was placed between them. Mott-Smith[6] used three Geiger-Müller counters placed in a vertical line and observed triple coincidences. The first two counters served as a "slit system" defining a beam which passed through the third. Between the second and third counters a block of iron is placed which can be magnetized. Observations were then made to determine the change in the number of triple coincidences when the iron was magnetized. The results were essentially negative, no change in the number of coincidences being observed.

This result may be interpreted as indicating that the particles have an energy of $2\times10^9$ electron-volts if electrons or half that amount if protons energies much higher than involved in the calculations on the assumption that the rays were electromagnetic waves. Work by Rossi indicates even higher energies. One way of accounting for the straight paths is to assume that the particles are neutron particles having zero charge and probably the same mass as the proton. Having zero charge they are not deflected by the magnetic field. In all the work with iron it is assumed that the field effective in deflecting a charged particle is equal to the magnetic induction $B$. If the effective field should be the magnetizing field $H$ no deflection would be expected. A test of whether $B$ or $H$ is to be used in such experiments might be made by sending swift beta-particles through a thin sheet of magnetized iron.

The visible paths and the discharges of the counters may be caused by particles of secondary origin, produced by a primary ultra-gamma ray. Such an explanation has some support in that some of the paths in the cloud expansion apparatus seem to be produced simultaneously and to have a common starting point. Even if these particles were neutrons they may have been produced by high energy photons. If neutrons are produced by bombarding atoms with alpha-particles (see preceding note in these columns) may not ultra-gamma-rays be able to do the same?

To accept the secondary origin for these rays one must reconcile the fact that the curves showing the reduction of number of the visible paths and coincident discharges of the G−M counters have the same form as those giving the ionization in closed vessels versus depth in water. Unless the range of the secondary rays produced by the ultra-gamma-rays is very short a different form of curve would be expected. The range of the particles however is not short. If the observed rays are secondary in nature one would expect their average energy to be less than that of the initial photon. But the energies mentioned just above are much higher than those indicated by the gamma ray absorption calculations.

In a short article published in the May first issue of the Physical Review Millikan and Anderson[7] describe experiments in which tracks 15 cm long are produced by means of a magnetic field of 17,000 to 20,000 gauss in air. Out of 1000 exposures, 34 show measurable cosmic-ray tracks and on 24 of these photographs tracks of positive particles are found. On 6 of the photographs associated tracks are found, four of these plates showing tracks of one positive and one negative particle. The energies calculated on the basis of the particles being electron and proton are found to range from 16 to 500 million electron-volts with possibly a few of them having energy as high as 1000 million electron-volts. These authors conclude that the primary cosmic rays are not neutrons but are photons with energies corresponding to the atom-building of oxygen, silicon and iron. They believe that most of the tracks observed are produced by the bombardment of the nuclei of the air molecules by these photons.

The data thus are contradictory, at least the interpretations are, and the question as

[6] Mott-Smith, Phys. Rev. **39,** 391 (1932).

[7] Millikan and Anderson, Phys. Rev. **40,** 325 (1932).

to the nature, energy and source of that which has been called cosmic rays becomes more puzzling than ever. Perhaps these rays are combinations of photons electrons neutrons and possibly protons and the various types of experiments bring to the attention of the observer one or more of the kind of particles or waves that are coming to the earth from the outer world.

J. W. BUCHTA

# IMPORTANT NOTICE TO AUTHORS AND SUBSCRIBERS

•

The American Institute of Physics has taken over the editorial mechanics and business management of the journal PHYSICS.

Please note carefully the addresses to which correspondence should be sent.

•

**JOHN T. TATE, *Editor***
**University of Minnesota**
**MINNEAPOLIS, MINNESOTA**

will continue to receive:

Manuscripts submitted for publication

All correspondence concerning manuscripts prior to their acceptance or rejection.

•

**THE EDITORIAL SECRETARY**
**American Institute of Physics**
**11 East 38th Street**
**New York City, New York**

will receive:

All correspondence concerning the details of publication of accepted articles

All proof of articles in process of publication

All correspondence concerning reprint orders

All subscriptions and renewals

All orders for back numbers

# Important Notice to Contributors

•

A JOURNAL which publishes the results of research renders a service not only to the reader but also to the author and to the institution which sponsors his work. No research in pure science may be regarded as complete until its results have been made a matter of public record. From that point of view the publication of research may be regarded as as much an integral part of the research as any other mechanical feature of it. Considerations of this kind have lead the American Physical Society, The Optical Society of America, The Acoustical Society of America and The Society of Rheology to adopt the general policy of levying a charge of $2.00 per page against the material published in their official journals of research.

In keeping with this policy a charge of $2.00 per page will be made for articles published in PHYSICS beginning with the issue of July. Bills for this amount will be sent to the authors with instructions to present them for payment to the institution supporting their work. In no case is the author expected to pay the charge himself. Should for any reason the institution be unable to make payment the bills should be returned marked "no funds available."

In partial return for this charge the journal PHYSICS will supply 100 unbound reprints of each article to the institution.